Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
Department of Psychology, University of Michigan
Ann Arbor, Michigan

DAVID A. GRANT, Associate Editor
University of Wisconsin

DELOS D. WICKENS, Associate Editor
Ohio State University

Consulting Editors

E. JAMES ARCHER, *University of Wisconsin*
JUDSON S. BROWN, *State University of Iowa*
CLETUS J. BURKE, *Indiana University*
PAUL M. FITTS, *Ohio State University*
FREDERICK C. FRICK, *Massachusetts Institute of Technology*
FRANK A. GELDARD, *University of Virginia*
JAMES J. GIBSON, *Cornell University*
CLARENCE H. GRAHAM, *Columbia University*
HAROLD W. HAKE, *University of Illinois*
LLOYD G. HUMPHREYS, *University of Illinois*
ARTHUR L. IRION, *Tulane University*
HOWARD H. KENDLER, *New York University*
DONALD B. LINDSLEY, *University of California, Los Angeles*
KENNETH MACCORQUODALE, *University of Minnesota*
QUINN MCNEMAR, *Stanford University*
L. STARLING REID, *University of Virginia*
KENNETH W. SPENCE, *State University of Iowa*

CONTENTS

Effects of Exposure Time and Intertrial Interval Upon Decrement to the Müller-Lyer Illusion: P. T. MOUNTJOY 97

Human Trial-and-Error Learning Under Joint Variation of Locus of Reward and Type of Pacing: C. E. NOBLE AND J. L. NOBLE 103

Conditioning and Hypnosis: J. P. DAS 110

Stimulus Similarity and Sequence of Stimulus Presentation in Paired-Associate Learning: E. Z. ROTHKOPF 114

Associative Transfer in Motor Paired-Associate Learning as a Function of Amount of First-Task Practice: C. C. SPIKER AND R. B. HOLTON 123

Recognition of Novel Visual Configurations With and Without Eye Movements: C. M. MOONEY 133

Verbal Conditioning and Extinction With Verbal and Nonverbal Reinforcers: A. H. BUSS, I. R. GERJUOY, AND J. ZUSMAN 139

Meaningfulness of Material, Distribution of Practice, and Serial-Position Curves: H. W. BRAUN AND S. P. HEYMANN 146

Acquisition and Extinction of a Partially Reinforced Running Response at a 24-Hour Intertrial Interval: S. WEINSTOCK 151

Effects of Shock-Induced Stress on Verbal Performance: W. D. CHILES 159

Anxiety Level and Ego Involvement as Factors in Concept Formation: C. V. ROMANOW 166

R-S Learning as a Function of Meaningfulness and Degree of S-R Learning: E. M. JANTZ AND B. J. UNDERWOOD 174

Comment on "Role of Prefeeding in an Apparent Frustration Effect": A. AMSEL 180

Supplementary Report: Stability of a Verbal Rate Change in Experimental Anxiety: F. H. KANFER 182

Supplementary Report: Complexity and Orienting Responses With Longer Exposures: D. E. BERLYNE 183

Erratum 184

American Psychological Association

Vol. 56, No. 2 August 1958

ARTHUR C. HOFFMAN, Managing Editor
HELEN ORR, Circulation Manager
SADIE J. DOYLE, Editorial Assistant

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc., at Prince and Lemon Sts., Lancaster, Pa. The subscription rate per volume is $8.00, or $16.00 annually. Single copies are $1.50. Subscriptions, orders, address changes, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This Journal publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects, except when the latter are specifically oriented toward the extension of behavior theory. Experimental psychometric studies and studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology.

Normally, articles exceeding 20 printed pages cannot be accepted. Within this limit, the piecemeal experiment-by-experiment reporting of psychological research is discouraged, and the reporting of the data and conclusions of substantial segments of research efforts is encouraged. Specifically, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) *must* be presented in a single article. Special provision is made for the publication of Supplementary Reports and Replication Reports in articles of not more than 1.5 printed pages (see Editor's Note, this Journal, 1957, 53, 1–2).

Address all articles submitted for publication to the Editor, Arthur W. Melton, Department of Psychology, University of Michigan, Ann Arbor, Michigan. **Manuscripts *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc., as described in the *Publication Manual of the American Psychological Association*.** When in doubt about practices of the Journal, authors should examine a recent issue.

Articles are published in the order of their receipt, except in rare circumstances. Authors are supplied with 50 free offprints without covers. All of the cost of an author's alterations in galley proof is charged to the author. Priority in publication is given to articles whose authors assume the full cost of publication. In 1958 the publication cost is $20 per printed page. Authors of priority publications receive no free offprints.

To save printing costs, supplementary material to articles in this Journal are deposited with the American Documentation Institute. Copies may be obtained by ordering the Document Number given in the footnote reference to ADI. Orders should be sent to ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. Orders *must* be accompanied by checks or money orders payable to the Chief, Photoduplication Service, Library of Congress.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879. Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. and R. of 1948, authorized September 11, 1947.

Send address changes to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must reach the Subscription Office by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

© 1958 by the American Psychological Association, Inc.

Journal of
Experimental Psychology

VOL. 56, NO. 2 AUGUST, 1958

EFFECTS OF EXPOSURE TIME AND INTERTRIAL INTERVAL UPON DECREMENT TO THE MÜLLER–LYER ILLUSION [1]

PAUL T. MOUNTJOY

Denison University

Decrement in magnitude of constant error to the Müller-Lyer figure was reported to occur as a function of trials early in this century (**3, 5, 6, 11**). Recently interest in this phenomenon has been revived (**7, 8, 9**). Köhler and Fishback (**7, 8**) reported: an inverse relationship between number of trials per day and magnitude of decrement; greater decrement with spaced trials; additional decrement during overnight rest periods. Preliminary research (**12**) failed to verify these relationships and suggested alternative hypotheses. Decrement in magnitude of a response as a function of trials has long been referred to in the field of learning as "habituation" (**4**). If the change in response to the Müller-Lyer figure is an habituatory decrement, it is related to experimental variables in the same manner as are habituatory decrements of other responses.

Massed trials produce greater habituatory decrement (**1, 14, 15**). However, the intertrial interval effect was not significant in a previous study of the illusion (**12**). Extraneous stimulation has been shown to retard habituation (**2, 13, 15**). Because the addition of an intertrial task results in movement-produced stimulation, the retarding effect of spaced trials is accentuated by the addition of an intertrial task. Conversely, massed trials and/or the omission of an intertrial task results in greater decrement. Increased duration of stimulation resulted in larger habituatory decrement (**16**). Spontaneous recovery of habituated responses has been reported (**2, 10, 13, 14, 15**) and a greater number of habituation trials resulted in less spontaneous recovery (**14, 15**). Habituatory decrements are negatively accelerated and decreasing functions of number of trials (**1, 2, 10, 13, 16**), until the asymptotic level of response has been reached. These find-

[1] This paper is based upon a dissertation submitted to the Graduate School of Indiana University in partial fulfillment of the requirements for the Ph.D. degree. The author is indebted to J. R. Kantor, W. K. Estes, and H. G. Yamaguchi for many helpful suggestions during the course of this study.

TABLE 1

EXPERIMENTAL DESIGN

Intertrial Interval (ITI)	Exposure Time (ET)		
	0 Sec.	20 Sec.	50 Sec.
5 sec.	ET 0 AP 5 ITI 5 Total 10 Groups 3A and 3B	ET 20 AP 5 ITI 5 Total 30 Groups 4A and 4B	ET 50 AP 5 ITI 5 Total 60 Groups 5A, 5B, and 5C
25 sec.	ET 0 AP 5 ITI 25 Total 30 Groups 2A and 2B		
55 sec.	ET 0 AP 5 ITI 55 Total 60 Groups 1A, 1B, 1C, and C1A		

Note:—Each cell is divided into its components of exposure time, adjustment period (AP), and intertrial interval.

ings allow analogous predictions to be made for decrements to the Müller-Lyer figure.

METHOD

Design.—Intertrial intervals and exposure times were varied in accordance with an inverted L design (Table 1). This basic design was twice replicated, and all *S*s were run on two consecutive days. The first replication (Groups 1A, 2A, 3A, 4A, and 5A) received 31 trials on Day 1. The second replication (Groups 1B, 2B, 3B, 4B, and 5B) received only 13 trials on Day 1. The procedure of Group 1A was replicated without the intertrial task in Group C1A. The procedures of Groups 1A (Group 1C) and 5A (Group 5C) were replicated with a control figure on Day 1. All *S*s received seven trials with the illusion figure on Day 2. There were 18 *S*s randomly assigned to each of the 13 groups.

Apparatus.[2]—The illusion figure was the Brentano figure with the connecting lines omitted. The three arrowheads were etched into sheets of clear Lucite with each arrowhead consisting of two lines 37 mm. long which met at a right angle. The two arrowheads of the standard stimulus were on a single sheet with their points 153 mm. apart. The comparison stimulus arrowhead was on a separate sheet and could be moved from an extreme "in" position which represented an illusion of 125 mm. to an extreme "out" position which represented a negative illusion of 29 mm. An interchangeable pair of Lucite sheets containing a control figure was similarly constructed. The control figure was three vertical lines, each 74 mm. high. The vision of *S* was confined to the central portions of the Lucite sheets by a frame with inside measurements of 402 mm. in width and 191 mm. in height. A constant distance of 774 mm. between *S* and the figures was maintained by a facepiece. Dark adaptation during the intertrial interval was prevented by a white fluorescent tube mounted to illuminate the space between the facepiece and the shutter used to control the exposure of the figures. The *S* adjusted the comparison stimulus by means of a knob located directly below the center of the facepiece, and the position of the comparison stimulus was automatically recorded by a paper chart drive. A telegraph key was mounted directly below the knob for the intertrial task. The mechanism which raised and lowered the shutter was controlled by an automatic timer, which also rang a bell 5 sec. before the shutter closed. The bell served as a signal for *S* to respond by adjusting the comparison stimulus. This signal coincided with the opening of the shutter for those groups which did not have a preadjustment exposure time.

Subjects.—Data were obtained from a total of 234 volunteers of both sexes from the introduc-

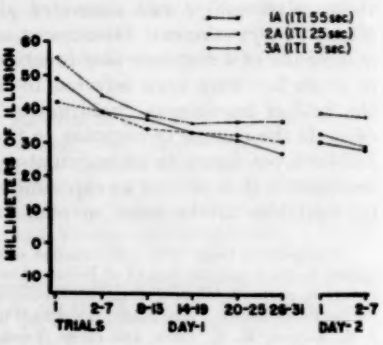

FIG. 1. Millimeters of illusion as a function of trials, days, and intertrial interval.

[2] A more detailed description of the apparatus, including a line drawing, may be obtained by ordering Publication No. 22,701 and remitting $2.00 for microfilm or $12.70 for paper enlargements from University Microfilms, 313 N. First St., Ann Arbor, Michigan.

tory psychology lecture sections at Indiana University. None of these volunteers wore glasses for any visual task except reading. An additional 42 *Ss* were run but were discarded because of violations of experimental conditions.

Procedure: Day 1.—During the reading of the instructions a measure of initial magnitude of illusion was obtained under identical conditions of a 5-sec. adjustment period for all groups in the guise of a practical trial. The *S* was instructed (with the aid of a line drawing) to set the point of the comparison arrowhead so that the distance between it and the point of the center arrowhead was equal to the distance between the points of the standard arrowheads. With the exception of those in Group C1A (for whom the key was covered) all *Ss* were instructed to tap the telegraph key once each second when the shutter was closed. All A-condition *Ss* received an additional 30 trials and all B-condition *Ss* received an additional 12 trials.

Procedure: Day 2.—The magnitude of illusion was obtained under identical conditions for all groups as before. An additional six trials were conducted under the assigned experimental conditions. The *Ss* were then interrogated concerning the illusion figure by means of a standard set of questions.

Results and Discussion

Intertrial-interval effects.—The hypothesis that decrement of the illusion is greatest under massed trials and least under spaced trials was evaluated by comparison of Groups 1A, 2A, and 3A. Inspection of Fig. 1 indicated that massed trials produced more decrement during the last half of Day 1, and this relationship was maintained during Day 2. The presence of the obtained difference on Trial 1 of Day 2 indicated that the differential effects of Day 1 persisted over a 24-hr. period because this trial was administered to all groups under identical conditions. The obtained difference, while not statistically significant (Table 2), indicated the desirability of further research because Köhler and Fishback (**7, 8**) concluded that spacing of trials produced greater decrement.

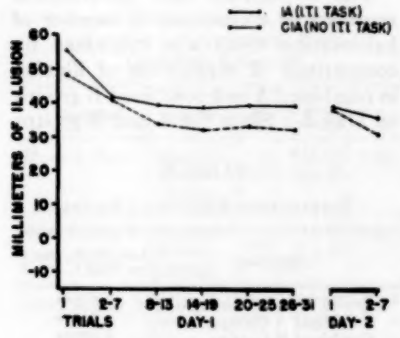

Fig. 2. Millimeters of illusion as a function of trials, days, and intertrial task.

TABLE 2

Trend Tests for Effects of Intertrial Interval, Intertrial Task, and Exposure Time

Variable	Source	*df*	*MS*	*F*
Intertrial interval (Groups 1A, 2A, and 3A)	Groups	2	1395.76	1.64
	Blocks	6	544.17	19.26**
	G × B	12	57.15	2.02*
Intertrial task (Groups 1A and C1A)	Groups	1	1404.86	1.62
	Blocks	6	212.54	3.17**
	G × B	6	55.63	0.83
Exposure time (Groups 3A, 4A, and 5A)	Groups	2	4403.93	4.14*
	Blocks	6	664.64	14.48**
	G × B	12	113.12	2.46*

* $P = .05$.
** $P = .01$.

The hypothesis that more decrement of the illusion occurs without an intertrial task was evaluated by comparison of Groups 1A and C1A. While the *F* value between groups did not attain significance (Table 2), the two groups were separated in the predicted order (Fig. 2). The presence of the obtained difference on Trial 1 of Day 2 indicated that the differential effects of Day 1 persisted over a 24-hr. period. Partial support of these two hypotheses indicated the desirability of further research upon the variables of intertrial interval and intertrial tasks.

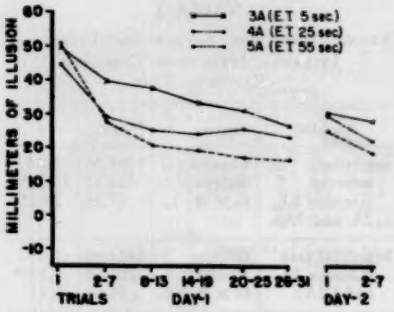

FIG. 3. Millimeters of illusion as a function of trials, days, and exposure time.

Exposure-time effects.—The hypothesis that decrement is greater with increasing periods of exposure to the figure was evaluated by comparison of Groups 3A, 4A, and 5A (Fig. 3). The trend test (Table 2) indicated that the null hypothesis could be rejected. The presence of the obtained difference on Trial 1 of Day 2 indicated that the exposure-time variable produced an effect which persisted over a 24-hr. period. Since increased exposure of the illusion figure produced more rapid decrement, the hypothesis was supported.

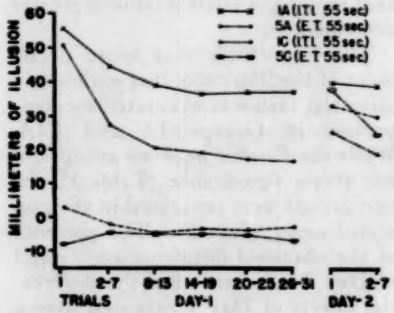

FIG. 4. Millimeters of illusion as a function of the Müller-Lyer figure and a control figure. Groups 1A and 5A set the illusion figure on Day 1 while Groups 1C and 5C set the control figure on Day 1. All groups set the illusion figure on Day 2.

Control-figure effects.—The procedures of Groups 1A and 5A were replicated in Groups 1C and 5C with the control figure on Day 1 and the illusion figure on Day 2 (Fig. 4). The different visual conditions had no effect on settings of the control figure. Although there was no decrement to the control figure on Day 1, significant decrement did occur to the Müller-Lyer figure on Day 2. The magnitude of the illusion for the control groups fell between the two experimental groups on Day 2. This indicated that part (but, not all) of the decrement in the experimental groups was due to learning not directly associated with the illusion figure.

24-hr.-interval effects.—The hypothesis that spontaneous recovery of the illusion occurs during a 24-hr. lapse of experimentation was evaluated by comparison of the last block of trials of Day 1 and Trial 1 of Day 2 (Table 3). Since inspection (Fig. 1, 2, 3, 4, and 5) indicated that spontaneous recovery occurred for all groups, this hypothesis was supported. This finding was contradictory to Köhler and Fishback's conclusion that further decrement occurred during a 24-hr. suspension of experimentation.

The hypothesis that spontaneous recovery is a function of number of habituation trials was evaluated by comparison of magnitude of illusion in combined A and combined B groups on Day 2. Since the A and B groups

TABLE 3

SPONTANEOUS RECOVERY *t* VALUES

Comparison	Last Block, Day 1 vs. Trial 1, Day 2
Pooled A and B Groups	5.938**
Combined A Groups	5.191**
Combined B Groups	3.079**

** $P = .01$.

were essentially identical before differential treatments began, and the B groups were (as predicted) statistically higher on Day 2 (Table 4), this hypothesis was supported.

The hypothesis that a larger number of trials results in greater decrement of the illusion has already been partially supported by significant *F* values for blocks of trials (Table 2). An additional test compared Block 2 and Block 5 of Day 1 for the combined A groups (Fig. 5). The *t* was significant far beyond the .01 level. This result was contradictory to Köhler and Fishback's conclusion that a large number of trials on a single day interfered with the occurrence of decrement.

Relationships between naiveté and decrement.—Köhler and Fishback reported that all their *S*s were naive, whereas, both naive and sophisticated *S*s were used in the present study. Any *S* who demonstrated appreciation of the illusory nature of the figure was classified as sophisticated. The occurrence or nonoccurrence of decrement for each *S* was determined for Day 1, Day 2, and the period between the two sessions. Since none of the three chi-square contingency tests were significant, these two variables were independent. This finding was in agreement with previous research (12) and indicated that discrepancies in the results of this study and that of Köhler and Fishback could not be attributed to sophistication of *S*s.

The contingency test between the two sessions threw further doubt on Köhler and Fishback's conclusion that additional decrement occurred in naive *S*s as a result of a 24-hr. suspension of experimentation.

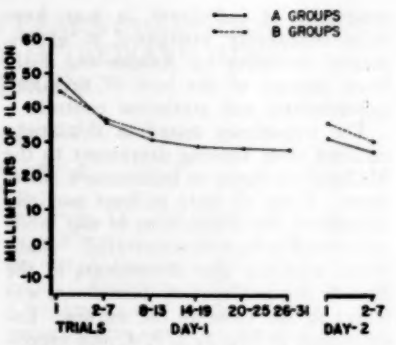

FIG. 5. Mean curves for the combined A and combined B groups. $N = 90$ for each curve.

TABLE 4

COVARIANCE ANALYSIS OF SPONTANEOUS RECOVERY OF COMBINED A AND COMBINED B GROUPS

Variable	Source	*df*	Reduced Variance	*F*
Block 2, Day 1	Methods	1	576.93	3.48
	Within Groups	177	165.57	
Trial 1, Day 2	Methods	1	1745.65	8.25**
	Within Groups	177	211.48	
Block 2, Day 2	Methods	1	1097.17	5.44*
	Within Groups	177	201.70	

* $P = .05$.
** $P = .01$.

The six hypotheses tested in this study were supported by the direction of obtained differences, and in four cases the differences were statistically significant. Because three hypotheses were contradictory to the conclusions of Köhler and Fishback, points of agreement must be summarized. (*a*) Statistically significant decrement did occur in naive *S*s even when no information was given concerning accuracy of response. This finding supported Köhler and Fishback's (7) criticism of the earlier interpretation that *S*s gradually learn to make their subjective judgments coincide with the objective world (17). (*b*) Negative settings of the illusion figure were confirmed in some *S*s. (*c*) Marked individual differences between *S*s were observed, and in part may account for discrepancies in results. Performances due to the idio-

syncrasies of individual *Ss* may have been mistakenly attributed to experimental variables by Köhler and Fishback because of the lack of necessary experimental and statistical controls.

The hypotheses tested in this study resulted from viewing decrement to the Müller-Lyer figure as habituatory decrement. Since all were at least partially supported, the fruitfulness of this interpretation has been demonstrated. Additional evidence that decrements in the illusion are habituatory decrements was found in the form of the curves. Examination of Fig. 1, 2, 3, 4, and especially 5 revealed that the decrements were of essentially the same form as habituation curves (**1, 2, 10, 13, 16**). These curves resembled learning and extinction curves in that the greatest changes occurred during the first few trials with the curves then becoming asymptotic. Continued research may disclose further similarities between decrements to the Müller-Lyer figure and other habituatory decrements.

Summary

A total of 234 undergraduate students was exposed to the Müller-Lyer illusion figure to determine the roles of experimental variables in the response decrements which occur during repeated trials. The principal findings were as follows:

1. Decrement in illusion magnitude occurred as a function of trials.

2. Decrement in illusion magnitude occurred more rapidly as exposure time to the figure increased.

3. Spontaneous recovery of the illusion occurred during the 24-hr. interval between experimental sessions.

4. There was an inverse relation between the number of trials on the first day and the magnitude of the recovered illusion at the beginning of the second day.

The results were interpreted as indicating that decrements to the Müller-Lyer illusion may be considered as a special case of the habituation paradigm.

References

1. Coombs, C. H. Adaptation of the galvanic response to auditory stimuli. *J. exp. Psychol.*, 1938, **22**, 244–268.
2. Davis, R. C. Modifications of the galvanic reflex by daily repetition of a stimulus. *J. exp. Psychol.*, 1934, **17**, 504–535.
3. Farnum, E. C. The Müller-Lyer illusion. In C. E. Seashore (Ed.), The effect of practice on normal illusions. *Psychol. Rev. Monogr. Suppl.*, 1908, **9**, No. 2 (Whole No. 38). Pp. 130–138.
4. Harris, J. D. Habituatory response decrement in the intact organism. *Psychol. Bull.*, 1943, **40**, 385–422.
5. Judd, C. H. Practice and its effects on the perception of illusions. *Psychol. Rev.*, 1902, **9**, 27–39.
6. Judd, C. H. The Müller-Lyer illusion. In C. H. Judd (Ed.), Yale psychological studies. *Psychol. Rev. Monogr. Suppl.*, 1905, **7**, No. 1 (Whole No. 29). Pp. 55–81.
7. Köhler, W., & Fishback, J. The destruction of the Müller-Lyer illusion in repeated trials: I. An examination of two theories. *J. exp. Psychol.*, 1950, **40**, 267–281.
8. Köhler, W., & Fishback, J. The destruction of the Müller-Lyer illusion in repeated trials: II. Satiation patterns and memory traces. *J. exp. Psychol.*, 1950, **40**, 398–410.
9. Köhler, W., & Wallach, H. Figural after-effects: An investigation of visual processes. *Proc. Amer. Phil. Soc.*, 1944, **88**, 269–357.
10. Lehner, G. F. J. A study of the extinction of unconditioned reflexes. *J. exp. Psychol.*, 1941, **29**, 435–456.
11. Lewis, E. O. The effect of practice on the perception of the Müller-Lyer illusion. *Brit. J. Psychol.*, 1908, **2**, 294–306.
12. Mountjoy, P. T. The effect of time of exposure and intertrial interval upon the rate of "destruction" of the Müller-Lyer illusion. Unpublished master's thesis, Indiana Univer., 1953.
13. Oldfield, R. C. Some recent experiments bearing on "internal inhibition." *Brit. J. Psychol.*, 1937, **28**, 28–42.
14. Piéron, H. Récherches expérimentales sur la phénomènes de mémoire. *Année psychol.*, 1913, **19**, 91–193.
15. Prosser, C. L., & Hunter, W. S. The extinction of startle responses and spinal reflexes in the white rat. *Amer. J. Physiol.*, 1936, **117**, 609–618.
16. Schoeffler, M. S. Decrement of leg withdrawal response in a dog as a function of shock duration. Unpublished master's thesis, Indiana Univer., 1953.
17. Woodworth, R. S. *Experimental psychology.* New York: Holt, 1938.

(Received October 28, 1957)

Journal of Experimental Psychology
Vol. 56, No. 2, 1958

HUMAN TRIAL-AND-ERROR LEARNING UNDER JOINT VARIATION OF LOCUS OF REWARD AND TYPE OF PACING[1]

CLYDE E. NOBLE AND JANET L. NOBLE

Montana State University

Previous experiments in this series on human trial-and-error learning have employed a nonverbal serial multiple-choice task called the Selective Mathometer (**6, 9**). Among the independent variables thus far investigated are the number of available responses (**5, 8**) and the number of choice points (**7**). As expected, difficulty is a positive function of both. All studies reported to date (**6**) have utilized a serial reinforcement procedure and an automatic *E*-paced rate of presentation. Since the apparatus was designed to permit the introduction of terminal reinforcement and an *S*-paced presentation rate as well, it is important to observe the effects of these altered conditions upon performance. The former will be referred to as the factor of *locus of reward*, the latter as the factor of *type of pacing*.

Aside from methodological considerations, an interesting theoretical issue is involved; i.e., the influence which the locus of reward has upon (*a*) rate of learning and (*b*) the shape of the serial position curve. According to Hull (**1, 2**), feeding a hungry animal at the conclusion of a series of acts sets up a gradient of generalized habit strength whose maximum occurs at the last choice point. On the other hand, if each of the acts is individually rewarded, separate gradients are generated. Assuming that these primary reinforcement gradients overlap and that some type of habit summation takes place: (*a*) Serial reinforcement ought to produce faster acquisition on the average than Terminal reinforcement, but (*b*) the serial error function in the latter case should rotate clockwise toward the posterior end of the chain.

This expectation is also consistent with a prediction from Hull's (**3**) delay-of-reinforcement postulate, the argument being that Terminal reward involves delayed reinforcement as compared with Serial reward, hence greater difficulty for all but the final act of the series. Spence (**10**) has recently suggested two alternative hypotheses whereby such differential performance may be explained. One involves a partial-reinforcement incentive-motivational interpretation of the classically conditioned r_g-s_g ("expectancy") mechanism. The other simply regards delayed-reward learning as analogous to experimental extinction, the latter constituting the limiting case of infinite delay in instrumental conditioning. To determine whether these formulations are applicable to human *S*s in secondary reinforcement situations, experimental manipulation of locus of reward is indicated.

The pacing factor is of interest because human *S*s are usually required to respond at a rate fixed by *E* or by apparatus,

[1] This report is based on work done under ARDC Project No. 7707, Task No. 77133, in support of the research and development program of the Air Force Personnel and Training Research Center, Lackland Air Force Base, Texas. Permission is granted for reproduction, translation, publication, use, and disposal in whole and in part by or for the United States Government.

D. E. Niemoeller collected the data in 1953, and analyses were completed under a grant from the Louisiana State University Council on Research. This phase was assisted by W. T. Alcock, V. J. Cieutat, F. J. Farese, R. L. Frye, Jr., D. A. McNeely, P. A. Ray, J. L. Regan, and R. E. Schluter. A portion of the paper was read before the Southern Society for Philosophy and Psychology, April, 1955.

whereas animal *S*s are nearly always self-paced (e.g., 1, 2). Although there is little reason to expect a difference between normal *E*-pacing rates and *S*-pacing, this must be confirmed before interspecies comparisons are possible. Pacing variations on the Mathometer do not change the basic task, but this would not necessarily be true of animal maze learning. Joint variation of type of pacing and locus of reward is necessary to discover whether these two factors interact.

The present experiment will investigate four-link noncorrection serial trial-and-error learning by massed practice under immediate Serial vs. Terminal reinforcement and under *E*-pacing vs. *S*-pacing. Of major interest will be (*a*) the locus of reward factor, (*b*) the pacing factor, (*c*) their possible interaction, (*d*) probability acquisition curves, (*e*) the Practice-by-Reward interaction, (*f*) latency data, and (*g*) serial position effects.

Method

Apparatus.—The Selective Mathometer, which has been described in detail elsewhere (6, 9), is a semi-automatic 19-key push-button apparatus. Problem stimuli are presented by electronically controlled slide projection, to which *S* responds by selecting keys in various arrangements on a semicircular panel. Secondary reinforcement (G) is provided by a green lamp centrally mounted below the projection screen, and both counter and polygraph records are taken of all stimulus and response events. In order to facilitate the measuring of latencies (R_t), the chart speed of the graphic recorder moved at 12 in. per min. for the two *S*-paced conditions. The *E*-paced conditions, as in all prior research, employed a rate of 6 in. per min., which is optimal for scoring responses (R) as initial corrects (lR+), corrects (R+), or incorrects (R−). In the present study lR+ = R+, so response probability $R_p = \Sigma R+/\Sigma R$.

Procedure.—A 2 × 2 factorial design (4) was used together with the four-choice task described earlier (5, 7). The factor of locus of reward compared the effects of (Serial) reinforcement following each correct choice with (Terminal) reinforcement following only the last of four correct choices of a trial. In either case, the delay and duration of reward were .00 and .75 sec., respectively. The factor of type of pacing compared the effects of stimuli presented regularly at a 3.44-sec. rate for a 2-sec. duration (*E*-pacing) with stimuli presented as rapidly as *S* responded (*S*-pacing). In either case, unlike earlier studies which employed the modified-correction procedure, *S* was permitted to press only *one* key at each choice point (noncorrection procedure). This was necessary for the two reward conditions to be comparable.

Another change involved the use of massed rather than distributed practice in order to maximize the similarity between the two pacing conditions. There was no intertrial interval beyond the usual 1.44-sec. interstimulus period, and the four stimuli followed each other in an invariant sequence. The slides contained verbal paralogs and the exposed keys were Nos. 4, 8, 12, and 16, as in corresponding conditions of the experiments on response availability (5) and task length (7).

The *S*s consisted of 96 basic airmen at Lackland Air Force Base, 24 of whom were assigned randomly to each of the four treatment combinations. The correct reaction sequence appeared in 12 different permutations for all groups, with two *S*s having a given permutation. Each *S*'s problem was to learn a sequence of four push-button responses to a criterion of five perfect trials. Following previous studies, *S* was instructed to use the forefinger of his preferred hand, to respond as quickly as possible at each choice point, to return to the fixation plate after each choice,[2] and to refrain from responding when the screen was dark. After the experiment, each *S* was isolated from the remaining pool of *S*s in order to prevent collusion.

Two points are worthy of emphasis. First, the reinforcing event (G) for the two *S*-paced groups was independent of stimulus offset; i.e., this was a case of self-pacing on any R, not necessarily R+. By contrast, Hull's experiments with rats (1, 2) employed the usual *S*-paced correction procedure which amounts to self-pacing on R+. Second, the reinforcer for the two Terminal groups was delivered on an all-or-none basis; e.g., if *S* failed to receive a green light following a given trial, he might have made one, two, or three R−'s. In other words, Terminal reward was delivered only on perfect trials.

Results

Trials to mastery.—The general results of the experiment are shown in

[2] This feature, corresponding to the centering partition in Hull's four-choice-four-link animal maze (3, p. 172), required *S* to initiate each act in the series from a comparable point in space.

TABLE 1

MEAN NUMBER OF TRIALS TO MASTERY FOR SERIAL VS. TERMINAL REWARD UNDER *E*-PACING AND *S*-PACING

Locus of Reward	Type of Pacing		Means
	E-Paced	*S*-Paced	
Serial	18.50	19.25	18.87
Terminal	45.04	43.42	44.23
Means	31.77	31.33	

TABLE 2

ANALYSIS OF VARIANCE OF TRIALS TO MASTERY

Source	*df*	*MS*	*F*
Locus of reward (G)	1	15,428.01	22.17*
Type of pacing (P)	1	4.59	<1.00
G × P	1	34.22	<1.00
Error (w)	92	696.03	
Total	95		

* $P < .001$.

Table 1, calculated as the mean number of trials required by each group to reach the criterion of five successive error-free sequences. Examination of the cell means indicates that Terminal reward results in much slower acquisition than Serial reward under both types of pacing, but that *E*-pacing is no different from *S*-pacing.

These conclusions are confirmed by the 2 × 2 factorial analysis of variance (**4**) summarized in Table 2. The main effect of locus of Reward (G) is very significant, but neither the Pacing (P) effect nor the Reward × Pacing interaction is significant. It is notable that giving reward terminally raised the trial means by a factor of 2.3 and the variances by a factor of 6.7. This evidence of heterogeneous variance was confirmed by Bartlett's test, which resulted in a χ^2 of 36.36 ($P < .001$).[3] Manipulating the locus of reward in human learning apparently has as powerful an influence on response variability as upon average proficiency.

Acquisition curves.—In view of the absence of a difference between *E*-pacing and *S*-pacing, the data for the pacing factor were pooled giving 48 *Ss* each in the Serial and Terminal groups. Figure 1 presents the acquisition curves under the two locus-of-reward conditions (G), plotted as the empirical probability (relative frequency) of a correct response (R_p) during successive five-trial blocks (N). In order to study the full range of acquisition, the curves have been extended to the point of mastery required by the slowest *S* (**5**, p. 95). Consistent with previous work on response availability (**5, 8**) and task length (**7**), both curves are asymmetrical and sigmoidal, conforming to the rational function

$$R_p = a(i)^{r^N} \qquad [1]$$

where a = asymptote of R_p (unity), i = initial probability ($\frac{1}{4}$), r = rate parameter depending on locus of reinforcement, and N = number of trials by the noncorrection procedure.

The best-fitting equations, obtained by minimizing the sums of squares of the residuals to determine the empirical parameter r, are as follows:

Serial G:

$$R_p = 1.00\ (.250)^{.872^N} \qquad [2]$$

Terminal G:

$$R_p = 1.00\ (.250)^{.967^N} \qquad [3]$$

The percentages of variance accounted for are 98.8% and 99.8% for Equations 2 and 3, and the inflection points

[3] Repeating the foregoing analyses nonparametrically with the Mann-Whitney test gave similar results.

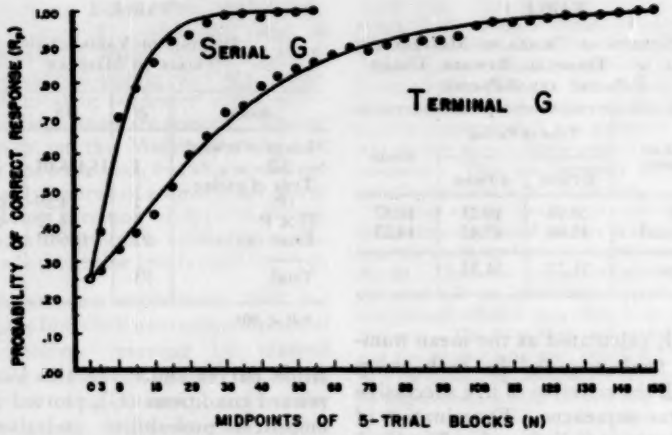

FIG. 1. Compound trial-and-error learning (Rp) as a function of practice (N) with locus of reward (G) as the parameter. The origin is rational.

occur at 2.5 and 12.0 trials, respectively. All calculational procedures have been described elsewhere (**7**) in greater detail. Further evidence is thus provided to support the hypothesis (**6**, p. 392) that Equation 1 is valid for trial-and-error learning phenomena, and specifically that r is governed by reinforcement contingencies (G) as well as by task factors (T).

A 7 × 2 Type I analysis of variance (**4**) was performed upon the R+ measure during the first seven 5-trial blocks for the Serial and Terminal groups. By restricting the range of possible scores as discussed in an earlier paper (**8**), extreme skewness and heterogeneity of form and variance were minimized. From the summary given in Table 3, it can be seen that both main effects and the interaction are highly significant. The Practice × Reward interaction, incidentally, confirms the difference in rate parameters and inflection points obtained in Fig. 1 and in Equations 2 and 3.

Latency data.—Due to the fact that the *E*-paced conditions provided no differential reinforcement for very fast reactions, and also because Table 2 failed to reveal any effect on R+ attributable to type of pacing, the analysis of R_t scores may be limited to correct choices made under the two *S*-paced conditions. Inspection of *E*-paced latencies showed no practice effect, but the *S*-paced latency curves indicated a slight superiority for the Terminal group during the final (asymptotic) stages of practice. However, the difference in mean R_t scores during Trials 31–55 was not significant (t = 1.48, 46 df). The same test repeated on speed scores ($R_{1/t}$) was also not significant (t = 1.32).

Pooling the two reward conditions gave a latency curve which began at about 1.50 sec. during Trials 1–5, dropped to about 1.25 sec. in the second block, then oscillated around 1.15 sec. for the remainder of the training period. There is, therefore, a systematic though not a differential increase in speed of reaction under the two *S*-paced reward conditions as a

TABLE 3

ANALYSIS OF VARIANCE OF CORRECT RESPONSES (R+) FOR SERIAL VS. TERMINAL REWARD DURING TRIALS 1–35

Source	*df*	*MS*	*F*
Between *Ss*	95		
Locus of reward (G)	1	7735.71	51.18*
Error (*b*)	94	151.15	
Within *Ss*	576		
Practice blocks (N)	6	1030.39	76.84*
N × G	6	114.08	8.51*
Error (*w*)	564	13.41	
Total	671		

* $P < .001$.

TABLE 4

SERIAL-POSITION EFFECTS IN TERMS OF MEAN LATENCIES IN SECONDS ($\bar{R}_t$) AND MEAN TOTAL CORRECTS ($\bar{R}+$) FOR SERIAL VS. TERMINAL REWARD UNDER THE *S*-PACED CONDITION DURING TRIALS 1–10

Score	*n*	Locus of Reward	Choice Point			
			I	II	III	IV
$\bar{R}_t$	23	Serial	1.46	1.35	1.41	1.29
	22	Terminal	1.75	1.36	1.25	1.13
$\bar{R}+$	24	Serial	5.46	5.63	6.00	5.27
	24	Terminal	3.04	2.75	2.58	2.63

function of amount of practice, the change becoming stabilized after about 30 trials.

Serial-position effects.—Unlike the foregoing analyses of R_t data by trials, the serial-position latency curve for the *S*-paced Terminal group during Trials 1–10 differs markedly from that of the *S*-paced Serial group. As shown in the upper half of Table 4, there is a regular decrease in the latency of correct choices from the first to the fourth choice point for the Terminal group. (The *n*'s are reduced due to three *Ss* who consistently erred at certain choice points.) A simple test of the observed trend is provided by computing a *t* ratio between the first and last serial positions. The ratio is very significant ($t = 4.88$, 21 *df*, $P < .001$). The corresponding test of the initial and final points for the Serial group is not significant ($t = 1.61$, 22 *df*, $P > .10$). The two groups have about the same average latency, but their trends appear to cross.

Similar analyses were performed on serial-position effects measured by R+ scores. In this case, presented in the lower half of Table 4, there was no crossing nor were the trends significant. This result is consistent with earlier observations (**5, 7, 8**) that frequency measures rarely show serial-position effects in human trial-and-error learning under these conditions. Why the latency measure is sensitive to the locus-of-reward factor with respect to intraserial phenomena but not to extended practice effects is not known. The R_t data are reminiscent of Hull's animal error curves (**2**, p. 19) in that Terminal reinforcement appears to shift difficulty from the posterior to the anterior portion of the chain. Apart from this resemblance, however, the quantitative aspects of the two sets of results are quite different. The classical bowed form fails to appear either in the R+ or in the R_t scores.

DISCUSSION

In the 1940's, Hull (**1, 2**) carried out two experiments on heterogeneous compound trial-and-error learning in rats using a quadruple-choice linear maze. He found that the first choice point was easiest under Serial reward but that the fourth choice point was easiest under Terminal reward; this observation agreed with his theoretical derivation from reinforcement and generalization principles. Hull apparently did not notice, however, that both groups originally required th

maze habit at about the same rate. Using his published data (1, 2) to calculate the mean percentage of initial-choice correct responses (1R+) cumulated during 50 distributed trials, one secures a value of 49.43% (σ^2_S = 52%) for the Serial condition (n = 48) and 53.74% (σ^2_T = 35%) for the Terminal condition (n = 72). Assuming a zero correlation the standard error of the difference is 9.3%; hence, the resulting critical ratio of .46 is not significant. Lacking any evidence of compensating differences in Hull's *S*s, apparatus, or procedure, this paradoxical result suggests that some kind of Serial reward must have occurred in the Terminal reward condition which equalized the two groups' proficiency.

One possibility is that by locking all but the correct door at a given choice point, Hull may have introduced secondary reinforcement by permitting locomotion in a goalward direction. Indeed, such a contingency is a necessary consequence of using the correction procedure in that maze (3, p. 192). With the Mathometer, however, this problem does not arise because correct and erroneous keys operate in exactly the same manner. Thus, in the present experiment the reinforcing event is not confounded with the aftereffects of responding.

There are three possible explanations for the effects of locus of reward: (*a*) Hull's habit-differential theory, (*b*) the Hull-Spence incentive-motivation theory, and (*c*) Spence's extinction theory. The first explanation (1, 2) appears unlikely because of the absence of serial-position phenomena which presumably are necessary concomitants of the facilitation resulting from Serial reward. Evidence for the second explanation (3, 10) is ambiguous; although Practice and Reward combine in a multiplicative fashion like habit strength and incentive motivation, it is uncertain whether human expectancy reactions in mathometry function like animals' fractional consummatory acts. The third explanation (10) seems to agree with the data and to be parsimonious as well. According to this counterconditioning view, Terminal reward leads to a greater probability of competing reactions occurring during a given trial than does the Serial reward condition. As in delayed-reward learning, uncontrolled activities fill the interval, thus interfering with the learning of the appropriate response sequence. The probability acquisition curves in Fig. 1 and the analysis of variance in Table 3 show that locus of reward affects merely the rate of acquisition, leaving both origin and asymptote unchanged. In terms of Equation 1, r varies with the reinforcement contingencies when a and i are held constant by the nature of the task.

Summary

This experiment was concerned with non-correction serial trial-and-error learning comparing the effects of Serial reinforcement with Terminal reinforcement, and of standard *E*-pacing with *S*-pacing. Ninety-six human *S*s, divided into four groups, learned a sequence of four push-button responses to a criterion of five perfect trials.

Terminal reward required a significantly greater number of trials to reach mastery, but there was no difference due to type of pacing and no interaction. Response probability acquisition curves (R_p) were about 99% in agreement with the rational equation $R_p = a(i)^{r^N}$, where a = asymptotic probability, i = initial probability, r = rate parameter governed by reward contingencies, and N = number of trials. There was a significant interaction between practice and reward factors, but both origin (i = .25) and final leyel (a = 1.00) were independent of locus of reward.

Although reaction latency decreased with training, there was no over-all difference due to locus of reward or type of pacing. As to serial-position phenomena, the only effect was a decrease in latency with proximity to Terminal reward, Serial reward producing no difference. Serial-position frequency scores were unaffected by locus of reward.

References

1. Hull, C. L. Reactively heterogeneous compound trial-and-error learning with distributed trials and terminal reinforcement. *J. exp. Psychol.*, 1947, **37**, 118–135.
2. Hull, C. L. Reactively heterogeneous compound trial-and-error learning with distributed trials and serial reinforcement. *J. exp. Psychol.*, 1948, **38**, 17–28.

3. Hull, C. L. *A behavior system.* New Haven: Yale Univer. Press, 1952.
4. Lindquist, E. F. *Design and analysis of experiments in psychology and education.* New York: Houghton Mifflin, 1953.
5. Noble, C. E. Compound trial-and-error learning as a function of response availability (N_R). *J. exp. Psychol.*, 1955, **49**, 93–96.
6. Noble, C. E. Human trial-and-error learning. *Psychol. Rep.*, 1957, **3**, 377–398 (Monogr. Suppl. 8).
7. Noble, C. E. The length-difficulty relationship in compound trial-and-error learning. *J. exp. Psychol.*, 1957, **54**, 246–252.
8. Noble, C. E. Supplementary report: Response availability in human trial-and-error learning. *J. exp. Psychol.*, 1957, **53**, 359–360.
9. Noble, C. E., & Farese, F. J. An apparatus for research in human selective learning. *J. Psychol.*, 1955, **39**, 475–484.
10. Spence, K. W. *Behavior theory and conditioning.* New Haven: Yale Univer. Press, 1956.

(Received May 13, 1957)

Journal of Experimental Psychology
Vol. 56, No. 2, 1958

CONDITIONING AND HYPNOSIS[1]

J. P. DAS

Institute of Psychiatry (Maudsley Hospital), University of London

Many investigators have suggested that hypnosis can be explained, at least partially, by a conditioned response mechanism (**8, 10, 13, 16, 18**). They believe, essentially, that hypnosis is a form of conditioning. However, none of them have demonstrated a positive relationship between hypnosis and conditioning. A study by Corn-Becker, Welch, and Fisichelli (**1**) comes nearest to an experimental support of the above belief. But it can only be inferred from their study that "abstract conditioning," as found by them, is an analogue of the hypnotic situation. It still remains to be shown, as Pattie (**12**) remarks, that *Ss* who acquire "abstract conditioning" are those more susceptible to hypnosis.

The present study is an attempt to discover the kind of relationship, if any, between conditioning and hypnosis.

METHOD

Measuring hypnosis.—Hypnosis was measured by a 7-point scale having an item reliability of .85 (obtained by using Kuder-Richardson Formula 20). A brief induction procedure of 15 min. duration was employed. It was suitable to induce a light to medium depth of hypnosis. Five hypnotic suggestions and two post-hypnotic suggestions were then given to test the depth of hypnosis attained by *S*. The seven suggestions were the usual ones used in testing hypnotic depth (**7**). An *S*'s score was based on the number of suggestions he could not reject. But the final assignment of a hypnotic score also considered the answers of *S* to an oral questionnaire given after hypnosis. The oral questionnaire included questions such as "Did you feel relaxed? Heavy? Sleepy? How did you feel when you were suggested that . . . ?"

The induction procedure, described at length by Das (**2**), was monotonous, consisting largely of the repetition of such words as "relax," "heavy," "drowsy," and "sleepy." The procedure was the same for all *Ss*.

Measuring conditioning.—Eyelid conditioning was employed as a measure of conditioning. The apparatus and conditioning variables have been described by Franks (**4, 5, 6**). Briefly, the UCS was a puff of air with a range of pressure between 60 and 65 mm. of mercury delivered to the right eye of *S* from a distance of approximately 2 cm. and was of 500 msec. duration. The UCS was preceded by the CS, a pure tone of 1110 cycles, 65 db above the auditory threshold of *S*, lasting for 800 msec. and preceding the UCS by 300 msec. Thus CS and UCS overlapped for 500 msec. and terminated simultaneously. Every *S* was given 30 reinforced trials interspersed with 18 nonreinforced trials (test trials) followed by 10 extinction trials. The total number of blinks during the test trials in the acquisition period was taken to be the score for acquisition and the number during the extinction period was regarded as the extinction score. The coefficient of reliability for the acquisition measure was .88 and for extinction .86, as estimated by an analysis-of-variance technique. An oral questionnaire was given to *S* soon after the conditioning experiment. The most important question was concerned with whether *S* detected a connection between UCS and CS, and his conscious reaction to the CS.

The actual aim of the conditioning experiment was not disclosed to *Ss* who were told that it was a test of relaxation; all they were supposed to do was to sit comfortably and keep their eyes open. Similarly, *Ss* were led to believe before hypnosis was attempted on them that they were to take part in a test of relaxation.

Subjects.—The *Ss* were 63 male English-speaking students, aged from 16 to 26 yr. They were paid volunteers. No significant correlations were obtained between age on the one hand and hypnosis, acquisition, and extinction on the other. Similarly, intelligence as measured by Nufferno Accuracy Score (**9**) had no significant correlation with any of the three variables.

Procedure.—The conditioning session followed the hypnotic one. At least 30 min. intervened between the end of the hypnotic session and the

[1] This paper is based on portions of a Ph.D. dissertation (**2**). The author is indebted to H. J. Eysenck and W. D. Furneaux for their help and guidance, and to A. R. Jensen for his critical reading of the manuscript.

beginning of the conditioning procedure. During this interval *Ss* filled out a personality questionnaire.

Hypnosis was attempted when *S* was comfortably lying on a sofa with the room partially darkened. The familiar method of eye fixation to a small bright light held about 18 in. above *S*'s forehead was adopted. The *Ss* were given repeated suggestions to relax. Suggestions of feeling heavy, drowsy, and sleepy followed in gradual stages. After 15 min. the depth of hypnosis attained by *S* was tested by giving the seven specific suggestions. The *Ss* were awakened by counting up to 15 in a firm voice.

The conditioning session was held in a soundproofed room; *S* was separated from *E* by a curtain. The *S* sat facing into a large cardboard booth which restricted his vision. He could not see *E* although *E* could see him.

The hypnosis session lasted for about 25 min. and the conditioning session for about 30 min.

Results

Product-moment correlations were computed between hypnotic scores and the scores for acquisition and extinction. The results are summarized in Table 1 Two sets of correlation coefficients are presented besides the one based on transformed scores The second set of correlations was obtained by excluding four cases because of defects in their conditioning records. Three were rejected because it was gathered from their answers to the oral questionnaire after conditioning that attitudes interfered with their conditioning. All the three *Ss* came to know about the real purpose of the test in the course of the experiment and reported having consciously resisted blinking in the nonreinforced trials. The fourth case was rejected because the records of his conditioned blinks were unusual. While in the records, the usual conditioned blink, like a spontaneous blink, is a sharp crest (although sometimes of a lower amplitude), those of the fourth *S* were very low amplitude flattened waves. Only once did he give the usual type of conditioned blink, and that during the extinction trials; this perhaps was a spontaneous blink falling by chance within the duration of the CS.

TABLE 1

Product–Moment Correlations Between Hypnosis and Eyelid Conditioning

Hypnosis and	*N*	*r*	*Z*	*p*
Acquisition[a]	63	.387	.409	.0014
Extinction[a]	63	.361	.378	.0034
Acquisition	59	.509	.669	<.0001
Extinction	59	.415	.442	.0005
Acquisition[b]	59	.486	.531	<.0001

[a] The product-moment correlation between acquisition and extinction is .697.
[b] Based on scores transformed by $\log (20 - x)$.

Since both the distributions of hypnotic scores as well as of eyeblink acquisition scores were positively skewed, the scores were transformed by $\log (20\text{-}x)$, which tended to normalize the distribution, thus permitting a more satisfactory estimate of *r*.

Discussion

A positive and moderately high correlation was obtained between hypnosis and conditioning (acquisition and extinction). This correlation may be interpreted as revealing a connection between hypnosis and conditioning. Of course, only eyeblink conditioning was used in the present study. While it is not certain that a general factor of conditionability exists (11), there seems to be no reason why hypnosis should be correlated only with eyeblink conditioning and not with other forms of conditioning.

However, it may be thought that, because in the present study conditioning followed hypnosis, the former facilitated the latter. It cannot be discovered whether this was so from the design of the present study. Even if half of the *Ss* had been given conditioning on the first occasion and the other half hypnosis on the first occasion, still the question could not be answered, for both might be

mutually facilitating. Since to obtain a correlation the two tests have to be given to the same individuals, it appears that the existence of mutual facilitation can neither be verified nor eliminated. Even if the observed positive relationship between hypnosis and conditioning were partly a function of mutual facilitation, the most probable conclusion is that some essential elements underlie the two. In other words, facilitation works by favorably affecting the mechanisms responsible for one through the other, so that a "practice effect" appears when one follows the other.

The other point concerns *S*'s attitudes about hypnosis and conditioning. It has been pointed out that attitudes interfere with conditioning. Human conditioning proceeds best when it comes close to being a kind of unconscious biological learning, as is presumably the case in the conditioning of lower animals (**14**). It has been seen that the rejection of the three cases whose conditioning was interfered with by conscious attitudes, improved the correlation between conditioning and hypnosis. From the answers to the oral questionnaire on conditioning, no other *S* appeared to have consciously perceived the connection between CS and the UCS. It is felt that, had *S*s been more successfully misinformed about the nature of the conditioning experiment, still higher correlations might have been obtained.

Similarly, past attitudes to hypnosis may affect an *S*'s hypnotic score (**3, 15, 17**). To the knowledge of the writer, none of the 63 *S*s discovered that the "test of relaxation" which they were told about was in fact, hypnosis. Thus, attitudes to hypnosis possibly held by the *S*s were probably irrelevant.

Thus, from the above considerations, including the fact that age and intelligence have no significant relationship with hypnosis and conditioning, it may be concluded that the correlation between hypnosis and conditioning is not a result of any obvious artifact. It indicates most probably an essential similarity between the two processes.

Summary

This experiment investigated the relationship between hypnosis and conditioning. Sixty-three male students were subjected to a 15-min. procedure suitable for the induction of a light to medium-deep hypnosis, and were given seven specific suggestions to test the depth of hypnosis attained. Hypnosis was measured by a 7-point scale based on the number of suggestions *S* could reject. Later, *S*s were tested for eyeblink conditioning where both acquisition and extinction scores were obtained. After rejecting four *S*s whose records were suspect, r between hypnosis and acquisition was .509 ($P < .001$) and r between hypnosis and extinction .415 ($P < .001$). It was concluded that the correlations indicate a basic similarity between hypnosis and conditioning.

References

1. Corn-Becker, F., Welch, L., & Fisichelli, V. Conditioning factors underlying hypnosis. *J. abnorm. soc. Psychol.*, 1949, **44**, 212–222.
2. Das, J. P. An experimental study of the relation between hypnosis, conditioning and reactive inhibition. Unpublished doctor's dissertation, Univer. of London, 1957.
3. Dorcus, R. M., Brintnall, A. K., & Case, H. W. Control experiments and their relation to theories of hypnotism. *J. gen. Psychol.*, 1941, **24**, 217–221.
4. Franks, C. M. An experimental study of conditioning as related to mental abnormality. Unpublished doctor's dissertation, Univer. of London, 1954.
5. Franks, C. M. Conditioning and personality: A study of normal and neurotic subjects. *J. abnorm. soc. Psychol.*, 1956, **52**, 143–150.
6. Franks, C. M., & Withers, W. C. R. Photo-electric recording of eyelid movements. *Amer. J. Psychol.*, 1955, **68**, 467–471.
7. Friedlander, J. W., & Sarbin, T. R. The depth of hypnosis. *J. abnorm. soc. Psychol.*, 1938, **33**, 453–475.
8. Furneaux, W. D. Hypnotic susceptibility as a function of working suggestibility. In L. M. LeCron (Ed.), *Experimental hypnosis*. New York: Macmillan, 1952.
9. Furneaux, W. D. *Manual of Nufferno Speed Tests*. National Found. for Educ. Res., London.
10. Hull, C. L. *Hypnosis and suggestibility, an experimental approach*. New York: Appleton-Century-Crofts, 1933.

11. JENKINS, J. J., & LYKKEN, D. T. Individual differences. *Annual Rev. Psychol.*, 1957, **8**, 79–112.
12. PATTIE, F. A. Theories of hypnosis. In R. M. Dorcus (Ed.), *Hypnosis and its therapeutic applications*. New York: McGraw-Hill, 1956. Pp. 1–30.
13. PAVLOV, I. P. *Lectures on conditioned reflexes*. Vol. 1. New York: Int. Publ., 1928.
14. RAZRAN, G. A direct laboratory comparison of Pavlovian conditioning and traditional associative learning. *J. abnorm. soc. Psychol.*, 1955, **51**, 649–652.
15. WEITZENHOFFER, A. M. *Hypnotism: An objective study in suggestibility*. New York: Wiley, 1953.
16. WELCH, L. A behavioristic explanation of the mechanisms of suggestion and hypnosis. *J. abnorm. soc. Psychol.*, 1947, **42**, 359–364.
17. WHITE, R. W. The prediction of hypnotic susceptibility from a knowledge of subject's attitude. *J. Psychol.*, 1937, **3**, 265.
18. YOUNG, P. C. Experimental hypnotism: A review. *Psychol. Bull.*, 1941, **38**, 92–101.

(Received September 6, 1957)

Journal of Experimental Psychology
Vol. 56, No. 2, 1958

STIMULUS SIMILARITY AND SEQUENCE OF STIMULUS PRESENTATION IN PAIRED-ASSOCIATE LEARNING [1,2]

ERNST Z. ROTHKOPF [3]

Rensselaer Polytechnic Institute

The sequence of presentation of similar stimulus terms during practice appears to be related to the rate at which a paired-associate list is learned. Gagne (1) found that when similar stimulus items were presented in immediate ordinal succession, *Ss* learned paired associates more rapidly than when exposed to practice sequences involving the maximal ordinal separation of similar stimuli. These findings were based on an experimental task which employed nonsense figures as stimulus terms and nonsense syllables as responses. Gagne's conclusions, if generally applicable, have important practical implications because they suggest that for many commonly trained paired-associate skills, successive presentation of similar stimulus material is a desirable practice method. The present experiment was intended to test the general applicability of these conclusions by systematically manipulating the sequence of stimulus presentation in an experimental paired-associate task different from that used by Gagne. The task selected for this purpose involved the use of aural Morse signals as stimuli and their written alphabetic or numeric equivalents as responses. Morse code signals were chosen as stimulus terms because: (*a*) signals of the International Morse Code include groups of signals which are very similar to each other; and (*b*) because alphabetic Morse code reception training is a task to which the Gagne findings were thought to be particularly applicable (1, p. 71).

Experiment I

The purpose of the first experiment was to test whether Gagne's generalization that the maximal clustering of similar stimuli speeds learning was applicable to the present paired-associate task.

Method

Materials.—The experimental list consisted of 12 paired associates. Every paired associate was composed of a single, aural Morse signal and its appropriate alphabetical or numerical equivalent. The 12 stimulus terms were composed of four groups, each of which consisted of three stimuli relatively similar to each other. These groups and their stimulus members, with the correct equivalent for each indicated in parentheses, were as follows: Group a: · – – (W), – – · (G), – · (R); Group b: · – – – – (1), · · – – – (2), · · · – – (3); Group c: · – · · (P), · · · · (F), · – · · (L); and Group d: · · · (S), · · · · (H), · · · · · (5). Average similarity among the stimuli within each group was 42.5%. Average similarity among stimuli which were not in the same group was 16.6%. Similarity as used here is the proportion of the total presentations in a psychophysical situation in which the two concerned stimuli are judged as "same" by Morse code-naive *Ss* (see 4). This measure has been shown to be related to the number of substi-

[1] This study was carried out in support of Project 7714 of the United States Air Force Personnel and Training Research Program.

[2] The author wishes to acknowledge the valuable assistance of Marvin Zelle who prepared and recorded the experimental tapes with great patience and skill; of Risden Westen who advised on the machine analysis of the data; and of Donald C. Howell who performed portions of the final data analysis. Special thanks is due to the staff of the Air Force Radio Operator School at Keesler Air Force Base, Mississippi, whose cooperation aided this study to an important degree.

[3] This study was made while the author was on the staff of the Maintenance Laboratory, Air Force Personnel and Training Research Center.

tution errors which occur between stimuli in the paired-associate situation (4) and is therefore related to the similarity criterion used by Gagne (1).

Experimental treatments.—The experiment compared two treatments—one involving the maximal separation, the other involving the maximal clustering of similar stimuli in practice. The methods for making up practice material was essentially the same as that used by Gagne in Cond. 1 and 2 and is described in somewhat greater detail in his paper (1). Briefly the experimental treatments were as follows:

The *Grouped Similar* treatment, henceforth referred to as Grouped-S, was characterized by the successive or clustered presentation of similar stimuli. The practice orders were arranged so that the stimuli of each of the groups listed above were always presented in sequence at every trial. For example, on the first trial the practice order was: F(c), P(c), L(c), R(a), W(a), G(a), 5(d), S(d), H(d), 2(b), 3(b), 1(b). The lowercase letters in parentheses indicate the group to which each stimulus belongs. The order of stimuli within groups, as well as the order in which the four groups appeared, was varied at random from trial to trial.

The *Separated* treatment was characterized by the maximal ordinal separation of the similar stimuli in the practice order. This was achieved by always interposing a stimulus from each of the three other subgroups between any two stimuli of a given subgroup. For example, the practice order on the first trial was: W(a), F(c), S(d), 3(b), G(a), P(c), H(d), 1(b), R(a), L(c), 5(d), 2(b). Again the over-all order was randomly varied from trial to trial, but three relatively dissimilar stimuli were always interposed between two similar stimulus terms.

Procedure and apparatus.—Each *S* received a total of 64 practice trials. Each trial consisted of a single presentation of each of the 12 paired associates of the list. A 12-item test run followed every fourth practice trial. Test runs consisted of a single random-order presentation of every one of the 12 Morse stimuli without the spoken equivalent.

Practice was conducted in four blocks which were separated by a 10-min. rest interval. Each block consisted of 16 practice trials and four suitably spaced tests which were recorded on a reel of magnetic tape. Every *S* was exposed to four different reels of practice tape. The details of the recording procedures were identical with those described elsewhere (3). It is only important to point out that each signal was sent fast enough (dot = .05 sec.) to make verbal analysis such as "that was three dots and a dash" unlikely. The *Ss* practiced by means of the code-voice method, a modification of paired-associate procedure, which is described in detail elsewhere (2). The version used here consisted of the presentation of a signal, which was followed by a 3-sec. silent period during which *S* could write his response. This period was in turn followed by the signal's spoken equivalent. The next stimulus followed 1.6 sec. later. Stimuli were presented every 4 sec. during tests.

The experiment was conducted in a well-lighted, windowless, air-conditioned room. Every seat was supplied with a pair of headphones which were connected by means of a separate line to one of two tape recorders. In this manner it was possible to administer two experimental treatments concurrently to a group of *Ss* sitting in the same room.

Subjects and experimental design.—All *Ss* were randomly selected from a population of students awaiting entrance into an Air Force school for jet engine mechanics. This school does not include radio code instruction in its curriculum nor does it require its students to have above-average Radio Operator Aptitude indices. The *Ss* who stated that they had previous Morse code training of any kind were rejected. The experimental design involved two replications (N = 18 and 24, respectively). In each replication both treatments were represented to an equal degree. The N for each experimental condition was 21.

Results and Discussion

The median percentage of correct responses per test for each experimental condition is shown in Fig. 1. The median was used in order to make this graph comparable with the results of Exp. III which is reported later in this paper. In all but the early stages of practice the Separated condition resulted in a greater number of correct responses than the Grouped-S treatment. An analysis of variance was performed to test the difference between the mean number of correct responses over all tests for the two experimental conditions. The resulting F was 4.98 ($P < .05$, 1 and 40 df).

It must be concluded that Gagne's generalization (1), that the clustering of similar stimuli in the practice order facilitates learning, did not hold for the present experimental task. However, the level of learning achieved in this experi-

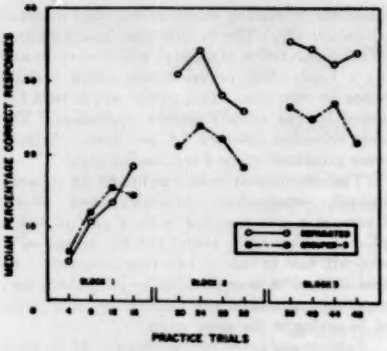

FIG. 1. Median percentage of correct responses as a function of practice for the two treatments of Exp. I. The median measure is shown in order to make this figure comparable to the results of Exp. III. The mean values for the data shown in this figure are in essential agreement with the median values.

ment was about 33% correct responses, while Gagne's *S*s attained about 75% correct responses. This limits the conclusions which may be drawn from the present experimental results because there is some suggestion in Gagne's data that early in learning the treatment which was equivalent to the Grouped-S condition resulted in a somewhat greater number of errors than the Separated condition.

EXPERIMENT II

The purpose of Exp. II was to determine whether the superiority of the Separated over the Grouped-S treatment is maintained even if practice is carried to a more advanced degree than that used in Exp. I. Experiment II also included two additional control groups which are described below.

Method

Experimental treatments.—Experiment II involved the Grouped-S and Separated treatments described in Exp. I and the following additional experimental conditions:

A Grouped, Relatively Dissimilar treatment, which will be referred to as Grouped-D, differed from the Grouped-S condition in that the sets of successively presented stimuli consisted of relatively dissimilar items. This was done in order to determine whether any effects associated with the Grouped-S treatment were due to the successive presentation of *similar* stimuli or resulted from the fact that three stimuli always appeared in ordinal conjunction (though not necessarily in the same order) from trial to trial. For this purpose the 12 stimuli were divided into four groups of three relatively dissimilar stimuli each (average within-group similarity: 16.2%). The members of each of these groups appeared consecutively at each trial. These stimulus groups were: W(a), 3(b), F(c); L(c), S(d), 1(b); P(c), G(a), H(d); and 2(b), 5(d), R(a). The foregoing was also the practice order used on the first trial. On subsequent trials W, 3, F, as well as each of the other groups were still together, but the within-group order and the order in which the groups appeared were changed. For example, the second trial order was: S(d), L(c), 1(b); 3(b), F(c), W(a); P(c), H(d), G(a); 2(b), R(a), 5(d).

In another condition, the Random treatment, the stimuli were presented in random order with the restriction that each stimulus appeared only once on each practice trial.

The two experimental treatments described above were respectively equivalent to Cond. 3 and 4 of the Gagne study.

Procedures and experimental design.—General procedures, apparatus, methods for recording practice and test materials, as well as the type of *S* used were the same as those in Exp. I.

Each *S* was exposed to a total of 10 blocks of practice, consisting of tapes appropriate to the experimental condition to which he was assigned for a total of 160 practice trials. A 20-hr. rest interval was interposed between Blocks 5 and 6. As in Exp. I, the rest interval between the other blocks was 10 min.

The experimental design involved eight groups of 10 *S*s each. Two such groups were assigned at random to each of the four experimental conditions. This design was balanced for time of day in that one group under each experimental condition was trained in the morning of two successive days while the other group was in training on two successive afternoons. Nine *S*s failed to return on the second day. These *S*s were not included in the analysis of data. The final *N*s for the four experimental treatments were as follows: Grouped-S—19; Separated—17, Grouped-D—18, Random—17.

Results and Discussion

In order to simplify graphical presentation, the scores for each *S* on the

four tests of any practice block have been combined. The median percentage of correct responses on the combined tests for each experimental condition is shown in Fig. 2. Again the median was used in order to make the data shown in this figure comparable to the data obtained from Exp. III. As may be seen in Fig. 2, the Separated treatment resulted in a greater number of correct responses than any of the other conditions. An analysis of variance which was performed on mean number of correct responses throughout the experiment under each treatment, resulted in an F of 2.79 ($P < .05$, 3 and 67 df). A t test between the Separated and the Grouped-S means was significant at the .01 level. Differences between the Separated and the Grouped-D, and between the Separated and the Random conditions, when evaluated by t tests, were significant at the .02 level.

Although there is some tendency for the acquisition curves of all four experimental conditions to diverge during the latter portions of Day 1 practice, none of the t comparisons among the Grouped-S, Grouped-D, and Random treatments were statistically significant.

The results of Exp. II confirm the findings of the first experiment that ordinal separation of similar signals in the practice order is a more effective training condition than ordinal clustering of similar signals. Experiment II also indicates that the Separated treatment is more conducive to learning of the present paired-associate task than either the Grouped-D or Random treatment.

It is also clear from both experiments that learning the 12 Morse code paired associates is a very difficult task. After 160 practice trials, the median percentage of correct responses for the most effective treatment of Exp. II is only just short of 50%. This may have been due in part to the relatively low aptitude of the present S population for learning Morse code and also to relatively massed conditions of practice.[4]

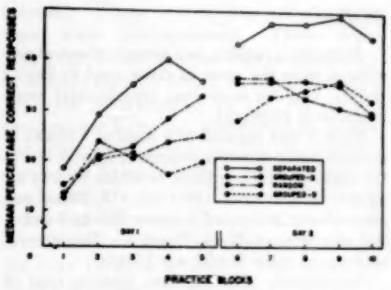

FIG. 2. Median percentage of correct responses as a function of practice in Exp. II. The median was used to make this graph comparable to Fig. 3. Each practice block is the equivalent of 16 practice trials. Data shown for each practice block is derived from four tests, one of which was administered after every fourth practice trial of the block.

EXPERIMENT III

The purpose of Exp. III was to replicate the conditions of Exp. II but to bring Ss to a greater degree of proficiency than had been possible in the previous two studies. This was done to determine whether the relative effectiveness of the four experimental treatments would be maintained for an appreciable portion of the training required for mastery. In order to accomplish this, Ss were selected from a population with relatively high general aptitude for Morse code training and more distributed conditions of practice were used.

[4] Examination of Fig. 1 and 2 suggests that progress in learning seems to be masked by the effects of massed practice. The acquisition data shown in these figures is characterized by little or no systematic progress during practice blocks. Marked improvement in performance, however, may be observed following rest intervals.

Method

Procedures, subjects, and design.—General procedures were the same as those used in Exp. I and II, and the same four experimental treatments were employed.

Each *S* was exposed to a total of 9 blocks of practice, using stimulus sequences appropriate to the experimental condition to which he was assigned, for a total of 144 trials. A 20-hr. rest interval was interposed between Blocks 3 and 4, and also between Blocks 6 and 7. The interval between all other blocks was 15 min.

Immediately after the last practice trial of each day, and also before the first trial of Day 2 and Day 3 of practice, a 48-item test was administered to all *S*s. This test consisted of four presentations of each of the 12 stimuli in random order. All test runs—those composed of 12 as well as those composed of 48 items—were identical for the four experimental treatments.

Apparatus arrangements similar to those in Exp. I were used. Every seat in the experimental chamber was supplied with a pair of headphones which were connected by a separate line to a switching console. The switching console permitted the selection of the output of one of four tape recorders for each pair of earphones. In this manner it was possible to administer four experimental treatments concurrently to a group of *S*s sitting in the same room. The switching console also permitted further amplification of the tape recorder outputs.

All *S*s were airmen awaiting entrance to an Air Force Radio Operator school who stated that they had no previous Morse code experience. All *S*s had a Radio Operator's Stanine of 6 or higher. It was planned to use four replications of 32 *S*s each. For administrative reasons as well as one case of equipment failure this plan could not be perfectly implemented. The *N*s for the four replications were, respectively, 30, 27, 29, and 21, for a total *N* of 107. Replication No. 4 did not include the Grouped-D treatment. In all other replications, all four experimental treatments were represented to an approximately equal degree. The *N*s for experimental treatments were as follows: Grouped-S—27, Separated—30, Grouped-D—21, and Random —29. The agreement between these four numbers and the *N*s for replications is coincidental.

Results and Discussion

The median percentage of correct responses for the tests of each practice block is shown for each experimental condition in Fig. 3. The Separated treatment produced a greater median number of correct responses than any of the other treatments throughout the course of the experiment. From the early practice sessions onward the distribution of correct response scores were too skewed to recommend analysis of variance. Instead, the Mann-Whitney *U* test was applied to the correct response scores for each of the three training days. For Day 1, the differences between the Separated treatment on the one hand, and the Grouped-S, the Grouped-D, as well as the Random conditions, respectively, were significant at the .05 level. For Day 2 the difference between the Separated and the Grouped-S condition reached the 5% level of significance.

Although the median number of correct responses made by *S*s under the Grouped-S treatment on either of the first two training days was somewhat less than that of *S*s operating under other conditions, the statistical comparisons did not support the conclusion that the Grouped-S, Grouped-D, and Random conditions consistently differed in training effectiveness. For Day 3 none of the comparisons among experimental con-

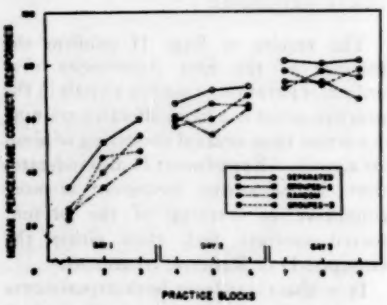

FIG. 3. Median percentage of correct responses as a function of practice in Exp. III. The median measure is shown because the statistical argument of the text did not utilize scores but rather the rank of scores.

ditions reached statistical significance.

The data obtained from the 48-item tests which were administered at the end of Day 1 and at the beginning and end of Days 2 and 3 were in essential agreement with the data reported above as far as the relative effectiveness of the four training conditions was concerned.

The findings of this experiment support those of the first two studies. It is noteworthy that the performance level achieved during the first practice block of Exp. III is about equal to that attained after the entire course of practice in Exp. I, while the performance level reached in Block 2 of the present experiment exceeds that attained after 10 blocks of practice in Exp. II. It is likely that the failure to obtain statistically significant differences among experimental conditions during the later stages of practice is at least in part due to the ceiling (144 correct responses) which the combined daily test scores could not exceed. The ceiling effect tended to reduce the usefulness of test measures as indices of learning effectiveness.

An attempt was made to determine whether differences among experimental treatments in correct responses were associated with differences in the incidence of any one of several categories of errors. Omissions constituted 7.44% of all errors while substitution errors accounted for 92.26%. Error responses, which were not appropriate to any of the stimuli used in this experiment, almost never occurred (.30%). There were no differences among treatments as far as relative frequency of omissions were concerned. Substitution error frequencies for the four conditions, however, were in approximately inverse order to the relationship among experimental treatments as far as correct responses were concerned.

Substitution errors were broken down further in order to evaluate the effects of repeated successive presentation of stimuli in the practice order on the frequency of substitutions among such stimuli. Three categories of substitutions were distinguished. These were: Type (a), substitutions within the groups of similar stimuli which were presented in repeated conjunction in the Grouped-S treatment; Type (b), substitutions within the arbitrarily chosen groups of relatively dissimilar stimuli which were presented conjunctively in the Grouped-D condition; and Type (c), all substitution errors which were neither Type (a) nor Type (b). As may be seen in Fig. 4, the mean frequencies of each of these three categories of errors generally were found to be in keeping with the relative standing of the four treatments as far as correct responses were concerned.

Discussion

The results of three essentially identical experiments have indicated that the Separated treatment facilitates learning more than the Grouped-S practice condition. This appears to hold for major portions of the practice required for mastery of the paired-associate skills under training. Data has also been presented from two experiments which show that the Separated treatment exceeds both the Random and the Grouped-D condition in training effectiveness. This could be demonstrated statistically for only the early portions of practice.

The four experimental treatments of this study differed in terms of two major related variables. These were: (*a*) the relative separation of similar stimuli and (*b*) the likelihood of certain stimuli appearing consecutively in the practice order. If the second of these variables had been effective in the present situation, it should have acted to delay the course of acquisition. This effect was expected because repeated conjunction of stimulus terms in practice can be presumed to set up serial habits which are likely to interfere with the paired-associate habits under training. The data do not support the hypothesis that the likelihood of certain stimuli appearing consecutively in the practice order is the only effective factor operating in the

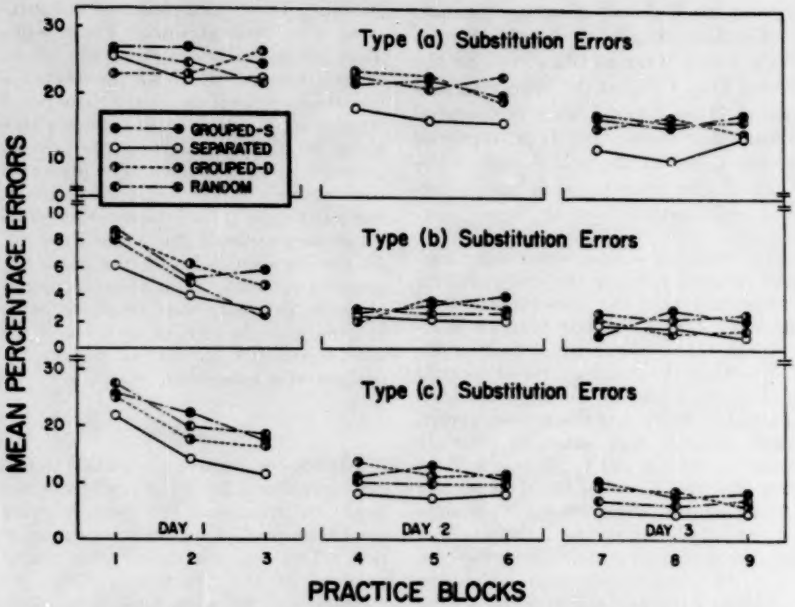

FIG. 4. Mean percentage of total responses of each of three categories of errors as a function of practice for the four treatments. The vertical scales of the three plots are not comparable. See the text for explanation of the error categories.

present experimental situation, because according to this hypothesis the experimental treatments should arrange themselves in descending order of training effectiveness as follows: 1. Random, 2. Separated, 3. Grouped-D, and 4. Grouped-S. This is due to the fact that repeated conjunction of certain stimulus terms is least likely for the Random condition, next likely for the Separated treatment, and by far the most likely for the Grouped-D and Grouped-S conditions.

The effect of repeated contiguity of certain stimulus terms can also be determined by examining the error data for a response tendency which will be called *anticipation set*. Anticipation set can be viewed as a serial habit which was thought to result from the fact that in the Grouped-S and Grouped-D conditions, three stimuli always appeared successively. For example, in the Grouped-S treatment, G, R, and W always appeared consecutively during practice. An instance of anticipation set here would be the increased likelihood of an R or W response to any stimulus which followed G. The term *practice mates* may be used to describe the members of any particular set of three stimuli which always appeared successively under one of the experimental treatments. Anticipation set is then defined as an increased tendency to give a response appropriate to a practice mate to *any* stimulus which follows a member of a set of practice mates.[5]

The data from the Grouped-S and the Grouped-D conditions were examined for antici-

[5] It can readily be seen that during randomly arranged test trials, such an anticipation set would tend to increase the number of errors if the stimuli which follow each other are not practice mates. The chances that any stimulus is not followed by a practice mate are greater than that it is followed by a practice mate in a random test trial. Therefore, the over-all effects of anticipation set should be an increase in the mean error scores of the Grouped-S and Grouped-D treatments.

pation set in the following manner. In every 48-item test, each stimulus was presented four times. In a few of the stimulus appearances the stimulus followed another practice mate in the order of test presentations. These will be referred to as AS—anticipation set stimuli. In most of the other appearances, the stimulus *did not* follow any of its practice mates. These will be referred to as FS—free stimuli. For every instance in which a particular stimulus was an AS, another instance was chosen at random in which the stimulus was an FS. Both of these instances were in all cases within the same 48-item test. The frequency of practice mate responses for each of these two presentations of the stimulus in point was then determined. For example, G followed the stimulus W at one point in the second 48-item test. The frequency with which G elicited G, R, or W in that position in the random test sequence was compared with the G, R, W frequency to G when that stimulus was at another arbitrarily chosen point (following the stimulus "5" in this particular example) in the second 48-item test. In the five 48-item tests, 23 such AS:FS pairings were found for stimuli appearing consecutively in the practice orders of the Grouped-S treatment. For these pairings, AS resulted in 503 practice mate responses, while FS gave rise to 511 such responses in *S*s trained under the Grouped-S treatment. The *S*s under the Random practice condition produced for the same AS:FS pairings, respectively, 558 and 567 practice mate responses.

A like comparison for stimuli appearing consecutively in the Grouped-D condition produced the following results. The *S*s under the Grouped-D treatment made 486 practice mate responses to AS, and 499 practice mate responses to FS. The corresponding response frequencies for the Random treatment were 697 for AS, and 714 for FS. These frequencies were based on 37 AS:FS pairings.

Since the number of practice mate responses to AS in no case exceeds like frequencies for FS, it must be concluded that no evidence for the interfering effects of anticipation set due to the Grouped-S or Grouped-D treatments could be shown for the present experiment.

Since the interfering effects of serial habits due to repeated ordinal conjunction of certain stimuli could not be demonstrated, the results of this experiment must be interpreted as supporting the notion that the maximal separation of similar stimuli during practice results in facilitated acquisition of paired-associate habits which involve Morse signals as stimulus terms. This study therefore does not support the generality for all paried-associate material of the findings previously reported by Gagne (1). Gagne found that the successive presentation of similar stimuli was the most effective practice sequence in a paired-associate task which employed nonsense figures as stimulus terms.

It is believed that these differences in findings are due to differences in the type of stimulus materials used in the two studies. Gagne's stimulus materials differed from those used in the present experiment in several respects. Among these is the frequency with which segments of the stimuli evoke distinct responses in experimentally naive *S*s. When presenting the Gagne nonsense figures, *E* can readily elicit from *S* such responses as: "This one has a sharp angle on top at the right"; "It has a long line sticking out in the lower left corner"; or, "The bottom of the figure is flattened"; and so on. These mediating responses, probably play an important role in the acquisition of such paired-associate material. In contrast, the Morse stimuli used in the present study were sent at a tone speed of 20 words/min. (standard word: Codez). This means that each individual dot is of .05 sec. duration. Such a tone speed in general does not facilitate analytic verbal responses to components of the signal. As a matter of fact, many *S*s are at a loss to describe signals (as by saying "it is a dot and four dashes") to which they have learned to respond with the appropriate alphabetic equivalent.

It is thought that the difference in stimulus material described above will submit to a simple rational account. This account cannot be made, however, without recourse to further experimentation. This is, to an important degree, due to the fact that the Gagne study differed from the present experiment in two other respects which may have been of consequence. These were:

1. Gagne presented stimulus and response terms simultaneously during practice trials. This resulted in a minimum of overt substitution errors during prac-

tice. The present experiment used the anticipation method. This resulted in many overt substitution errors during practice.

2. The response terms used by Gagne were trigrams consisting of a vowel placed between two consonants. Such trigrams are in general not perfectly in *S*s' response repertoire at the beginning of practice. This resulted in the fact that more than 60% of the total errors reported by Gagne were not strictly appropriate to any of the stimuli of the list. The response terms used in the present study, i.e., single letters or numerals, were very familiar to all *S*s prior to the beginning of practice. This assured the fact that nearly all errors were substitution errors.

Summary

Three experiments were performed which were designed to determine the role of sequence of stimulus presentation in the acquisition of paired-associate habits. The experimental task employed 12 aural Morse code signals as stimulus terms, and their appropriate alphabetic or numeric equivalents as response terms.

The first experiment ($N = 42$) compared practice orders (*a*) involving the maximal separation of similar stimuli with (*b*) orders involving maximal clustering of similar stimuli. Experiment II ($N = 71$) and Exp. III ($N = 107$) involved the comparison of four different methods of arranging practice orders. These were: (*a*) maximal separation of similar stimuli; (*b*) successive presentation of similar stimuli; (*c*) successive presentation of relatively dissimilar stimuli; and (*d*) random stimulus presentation.

The data indicate that the maximal separation of similar stimuli in the practice order led to facilitated acquisition of the paired-associate habits under training.

These results are contrasted with those of a previous study by Gagne (1).

References

1. Gagne, R. M. The effect of sequence of presentation of similar items on the learning of paired associates. *J. exp. Psychol.*, 1950, **40**, 61–73.
2. Keller, F. S. Studies in International Morse Code. I. A new method of teaching code reception. *J. appl. Psychol.*, 1943, **27**, 407–415.
3. Rothkopf, E. Z. A deduction from an excitation-inhibition account of retroactive inhibition. *J. exp. Psychol.*, 1957, **53**, 207–213.
4. Rothkopf, E. Z. A measure of stimulus similarity and errors in some paired-associate learning tasks. *J. exp. Psychol.*, 1957, **53**, 94–101.

(Received September 14, 1957)

Journal of Experimental Psychology
Vol. 56, No. 2, 1958

ASSOCIATIVE TRANSFER IN MOTOR PAIRED-ASSOCIATE LEARNING AS A FUNCTION OF AMOUNT OF FIRST-TASK PRACTICE

CHARLES C. SPIKER AND RUTH B. HOLTON[1]

Iowa Child Welfare Research Station

The classical design for associative interference in paired-associate learning is the A-B, A-C sequence. The *S* first learns one set of responses to a set of stimuli and then learns a new set of responses to these same stimuli. Studies of associative interference in verbal paired-associate learning have differed in the type of control condition against which the amount and direction of the associative transfer in the A-B, A-C series was assessed. One of the more common designs varies the amount of first-task practice given to different groups and compares the performance of these groups on the second task. Whether the amount of first-task practice is none versus some (**7**, **8**, **10**, **11**) or is varied through several degrees (**1**, **12**, **13**), the results have consistently showed over-all positive transfer, with relatively small amounts of interference restricted to the early trials.

Associative transfer has also been studied with unrelated lists. It has been established that positive transfer occurs from initial to subsequent unrelated lists (**4**, **14**, **15**) and that the amount of positive transfer is a direct function of the amount of immediately prior practice (**9**) and an inverse function of the time between first and second tasks (**5**). Thus, if one wishes to study the second-task interference resulting from the necessity for unlearning the responses learned to the stimuli in the first task (associative interference), the total amount of prior practice must be either statistically or experimentally controlled. The statistical control may be obtained by providing, at each level of first-task practice, a control condition in which *Ss* have had equivalent amounts of practice on a first task in which the associations are unrelated to those of the second task. Thus, Gibson (**3**) found that *Ss* given the A-B, D-C series performed significantly better on their second task than *Ss* given the A-B, A-C sequence. Since Gibson was not primarily interested in associative transfer, however, she neither equated the A-C and D-C lists nor counterbalanced them for difficulty, with the consequence that her results, in this respect, cannot be unequivocally interpreted. A similar reservation holds for the nonsignficant findings in the same direction by Bruce (**2**). Experimental control of prior practice may be obtained by giving all *Ss* the same amount of total practice prior to the transfer task, at the same time varying the proportion of total practice which is given on a task having the same stimulus items as those of the transfer task. The writers know of no experiments with associative interference in which

[1] From October, 1956, to May, 1957, the investigators collaborated in the design and conduct of a series of studies on associative transfer. The second investigator met her untimely death on May 9, 1957, in an automobile accident enroute to the school system in which the data were collected. The first investigator is responsible for the preparation of the manuscript. He is indebted to Don Lewis, I. E. Farber, and Boyd R. McCandless for helpful suggestions in its preparation.

amount of prior practice was experimentally controlled.

Four experiments were conducted to investigate associative interference in motor paired-associate learning. In Exp. I and II, amount of prior practice was statistically controlled; in Exp. III and IV, amount of prior practice was experimentally controlled. All experiments utilize intra-*S* comparisons in order to increase the precision of the experiments.

Experiment I

Method

Apparatus.—The apparatus was designed to present motor paired-associate learning tasks in which *S* learns to associate a different push button with each of six stimulus lights differing in hue. It consists of four parts: a stimulus presentation unit, the response unit, a control box, and a timing unit. It permits the presentation of a stimulus for a set anticipation time interval, followed by a timed interval during which the stimulus is presented jointly with an informational light beside the correct button, followed in turn by a timed interstimulus interval. The *S* is instructed to respond by pushing the correct button prior to the onset of the informational light. A bell sounds immediately if his response is correct, and, whether or not it is correct, the informational lamp is automatically turned on at the end of the anticipation interval.

The *stimulus unit* is a black box 4 in. high, 20 in. wide, and 12 in. deep. On the front face of the box is the stimulus aperture, 2 in. in diameter, covered with flashed opal glass. Inside the box, each of six 75-w. projection lamps is enclosed in a compartment on the front face of which is an opening over which a 2 × 2-in. glass color filter may be placed. Glass color filters obtained from the Corning Glass Co. were used. The stimuli were a red (No. 2404), a green (No. 4060), a yellow (No. 3387), a violet (No. 5113), an aqua (No. 4305), and a white (no filter). The stimulus box is mounted on 5-in. legs.

The *response unit* consists of a 12 × 20-in. black panel that slopes downward toward *S*. Arranged in a semicircle on the panel are nine push-button switches approximately 2½ in. apart. Beside each button, a small 6-v. pilot (informational) lamp is mounted with the glazed tip of the bulb projecting through the top surface of the panel. At the end of each anticipation interval, one of these lamps is energized to indicate which button *S* should have pushed. The buttons are numbered, starting at the left, from 1 to 9 with small white numerals painted on the panel beside the buttons. During the task, the response unit is placed directly in front of the stimulus unit so that the stimulus aperture is approximately 3 in. above the center button.

The *control unit* enables *E* to select the particular stimulus-response pairings appropriate for a given task and to determine which stimulus will be presented on a given trial. The *timing unit* consists of three Hunter decade interval timers to control the time between the onset of the stimulus and the onset of the informational lamp, the length of the joint presentation of the stimulus and the informational lamp, and the interstimulus interval. During the interstimulus interval, *E* turns a selector switch to that one of six positions which simultaneously determines for the next presentation the appropriate stimulus and the corresponding correct response button and informational lamp.

Subjects and design.—The *Ss* were 18 Grade 4 and Grade 5 children from the Ottumwa Public Schools.[2] All *Ss* were given three stimuli in Task I and were required to learn to associate a different one of the nine buttons with each stimulus. In Task II, they were given these same three stimuli and three stimuli they had not previously seen, and were required to learn to associate new buttons with each of the six stimuli. Thus, in Task II, half the light-button pairs represent an A-B, A-C paradigm and the other half, an A-B, D-C paradigm. Three *Ss* were randomly assigned to each of six subgroups in order to control for the possible differential difficulty of the stimulus-response pairs. The complete design is schematized in Table 1. The numbers in the table designate the buttons (from left to right) which were correct for each of the six color stimuli for Task I and Task II within each of the six counterbalancing subgroups.

Procedure.—The *S* was seated before the apparatus, shown the six stimulus lights one at a time, and asked to name them. He was then told that the experiment was concerned with a

[2] The investigators are indebted to R. O. Wright, Director of Curriculum, Ottumwa Public School System, Ottumwa, Iowa, for his generous permission to use the *Ss* of these experiments. They are also indebted to M. E. Wilson, Principal, Agassiz Elementary School, and his teaching staff for their fine cooperation in providing facilities and for their assistance in many other ways during the conduct of Exp. I and II. Thanks are also due Albert Whittlesey, Principal, Wilson Elementary School, and his teaching staff for equally fine cooperation during the conduct of Exp. III and IV.

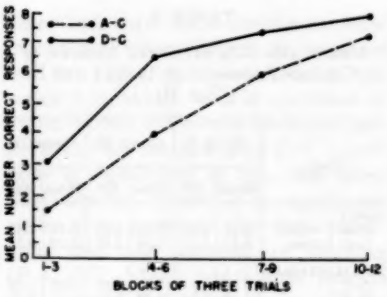

FIG. 1. Mean number of correct responses in the transfer task of Exp. I for both conditions.

comparison of adults and children in their abilities to discriminate among and to remember colors. He was given instructions concerning the nature of the task and was urged to guess on the first trial. The need to respond prior to the onset of the informational lights was emphasized. Immediately following the instructions, the first stimulus was presented and *S* was asked to respond.

During Task I, the anticipation interval was set at 3 sec., the joint presentation interval was 2 sec., and the interpresentation interval was 3 sec. Task I training for each *S* was continued until he met a criterion of three successive errorless trials, although all *S*s received at least 12 trials, where a trial was defined as the presentation of all three stimuli in some order. In each successive block of six trials, all six possible orders of presenting the three stimuli were used.

When *S* reached the Task I criterion, *E* announced that he was going to make the task harder by including more lights. He then set up the light-button connections which were appropriate for that *S*'s Task II. This required, on the average, about 2 min. No instructions were given *S* concerning the change in the light-button pairings from Task I to Task II. Next, *E* immediately began to administer 12 trials on Task II, where each trial included one presentation of each of the six stimuli. Six different orders of presenting the stimuli within a trial were used in each successive block of six trials. All three time intervals were set at 2 sec. for Task II.

The *E* recorded, on prepared sheets, the first button to which *S* responded on each stimulus presentation. Responses made by *S* to correct initial errors were recorded, but were not counted as correct in the tabulation of the data. Responses occurring after the onset of the informational light were recorded with a notation of their tardiness, but such responses were always scored as omissions in the analysis of the data.

TABLE 1

LIGHT–BUTTON CONNECTIONS IN TASKS I AND II FOR THE COUNTERBALANCING SUBGROUPS (Exp. I and Exp. II)

Stimuli	Subgroups											
	1		2		3		4		5		6	
	I	II	I	II	I	II	I	II	I	II	I	II
Red		5		4		6	2	9	3	8	1	7
Green		8		7		9	4	6	6	1	5	2
Yellow		1		2		3	7	3	9	5	8	4
Violet	2	9	3	8	1	7		5		4		6
White	4	6	6	1	5	2		8		7		9
Aqua	7	3	9	5	8	4		1		2		3

Results

The mean number of trials required to reach the criterion of three consecutive errorless trials in Task I was 6.33 ($SD = 3.50$). Since all *S*s were given a minimum of 12 trials, the majority of *S*s received considerable overtraining.

The performance of *S*s in Task II on the types of S-R connections is shown in Fig. 1. An analysis of variance of these data was conducted (**6**, Type VI, p. 292), with conditions (type of pair) and trial blocks as intra-*S* factors and the counterblanacing subgroups as an inter-*S* factor. A summary of this analysis is given in Table 2. Clear evidence is provided for the superiority of the D-C condition ($P < .001$) and a significant interaction ($P < .025$) was obtained between the Conditions and the Trial Blocks. The counterbalancing Subgroups effect was not statistically significant. The results of Exp. I thus permit the conclusion that the A-B, A-C series results in associative interference relative to the A-B, D-C.

EXPERIMENT II

Experiment II was conducted to determine whether or not the amount of interference in the

TABLE 2

ANALYSIS OF VARIANCE OF NUMBER OF CORRECT RESPONSES PER THREE-TRIAL BLOCKS FOR A-C AND D-C CONDITIONS (Exp. I)

Source	*df*	*MS*	*F*	*P*
Between *Ss*	17			
Subgroups (Sg)	5	11.58	1.21	
Error (b)	12	9.56		
Within *Ss*	126			
Trials (T)	3	196.80	102.50	.001
Conditions (C)	1	93.45	43.67	.001
T × Sg	15	2.35	1.22	
T × C	3	5.37	4.01	.025
Sg × C	5	3.78	1.77	
Sg × C × T	15	1.03	<1.00	
Error (w)	84			
Error$_1$ (w)[a]	36	1.92		
Error$_2$ (w)[b]	12	2.14		
Error$_3$ (w)[c]	36	1.34		

[a] Error term for T and T × Sg.
[b] Error term for C and C × Sg.
[c] Error term for T × C and T × C × Sg.

A-B, A-C condition is a function of the amount of first-task practice. Fifty-four children from Grades 5 and 6 were randomly assigned in equal numbers to three experimental groups.[3] Groups 5, 10, and 20 were given 5, 10, and 20 trials, respectively, on Task I. All *Ss* were then given 12 trials on Task II. All other procedures were the same as those in Exp. I.

Results

The means and *SD*'s of total correct responses for the three groups in different phases of the experiment are shown in Table 3. Correct responses for the first five trials of Task I indicate that the three groups were initially rather well matched, a simple analysis of variance yielding an *F* ratio less than unity. A similar analysis of the number of correct responses in the last five trials for the three groups yielded a significant *F* ratio ($P < .01$). The mean of Group 5 differed significantly ($P < .01$) from those of Groups 10 and 20; the latter means did not differ significantly. The different amounts of first-task practice were thus adequate to bring about different performances between Group 5 and the other two groups.

[3] See Footnote 2.

TABLE 3

MEANS AND *SD*'s OF TOTAL NUMBER OF CORRECT RESPONSES IN TASKS I AND II (Exp. II)

Task	Group 5		Group 10		Group 20	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
Task I						
First 5 trials[a]	6.17	3.20	7.22	2.69	6.11	2.85
Last 5 trials	6.17	3.20	13.11	2.05	14.77	0.55
Task II (12 trials)						
A–C	14.33	6.84	17.78	5.81	15.61	7.18
D–C	17.28	5.87	22.28	5.93	23.22	4.82

[a] For Group 5, the first 5 trials were also the last 5 trials.

The mean numbers of correct responses per three-trial block are shown in Fig. 2. Here it may be observed that the difference between the A-C and D-C curves increases with increasing amounts of first-task practice. A preliminary analysis of these data indicated that the subgroups did not differ significantly from one another and that they did not interact significantly with other variables. Consequently, a simpler analysis (6, Type I, p. 267) was conducted and a summary is presented in Table 4. For this analysis, only Groups 5 and 20

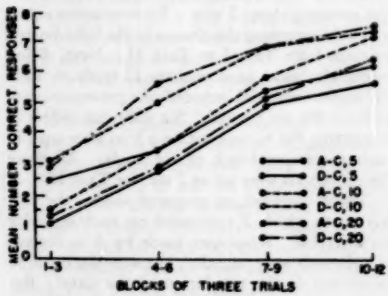

FIG. 2. Mean number of correct responses in the transfer task of Exp. II for all three groups under each condition.

were used, since these are the extreme groups and permit the most sensitive test for the trend predicted by the major hypothesis. As is apparent in Table 4, the difference between performance on the A-C and D-C conditions is significant at the .001 level. Furthermore, there is a significant interaction between the two conditions and the amount of training ($P < .05$). This interaction reflects the greater interference for Group 20 relative to Group 5.

Related t tests were used to compare the A-C and D-C means separately for each group. Performance on the D-C condition was significantly superior to that on the A-C for each group (Group 5, $P < .05$; Group 10, $P < .025$; and Group 20, $P < .001$). Thus, even five trials on Task I were adequate to produce statistically significant interference on the A-C condition. A simple analysis of variance of the number of correct responses for the three groups on the D-C condition alone showed that performance on the D-C condition improved significantly ($P < .005$) with increasing first-task practice. On the other hand, the same analysis for the A-C condition showed no significant variation in A-C means with first-task practice ($P > .10$).

TABLE 4

ANALYSIS OF VARIANCE OF CORRECT RESPONSES FOR GROUPS 5 AND 20 (Exp. II)

Source	*df*	*MS*	*F*
Between *Ss*	35		
Groups (G)	1	220.5	3.49
Error (b)	34	63.1	
Within *Ss*	36		
Conditions (C)	1	480.5	24.64**
C × G	1	88.9	4.56*
Error (w)	34	19.5	

* $P < .05$.
** $P < .001$.

TABLE 5

TOTAL FREQUENCY OF DIFFERENT TYPES OF ERROR ON TASK II (Exp. II)

Type of Error	Group 5		Group 10		Group 20	
	A-C	D-C	A-C	D-C	A-C	D-C
Omissions	98	102	108	58	105	70
Intralist errors	169	168	148	137	175	116
Task I intrusions	123	67	72	51	87	48
Old habit errors	(41)		(39)		(51)	
Total errors	390	337	328	246	367	234

The interpretation of these findings assumes that variation in amount of first-task practice simultaneously varies both positive and negative transfer. The positive transfer is presumably the result of nonspecific factors such as warming-up (**5, 9**) and/or learning-to-le rn (**4, 14, 15**). The negative transfer is the result of learning specific incompatible associations in Task I. Performance on the D-C pairs reflects only the positive transfer effects. Performance on the A-C pairs reflects both the positive and negative transfer effects. These opposing effects resulted in no significant net transfer for the A-C condition. The evidence for associative interference is obtained by comparing *Ss*' performance under the A-C and D-C conditions, in which comparison the positive transfer is held constant.

It is informative to consider the frequency of different types of error during the learning of Task II. The total frequencies are shown in Table 5, where "omissions" are defined as failures to respond during the anticipation interval, an "intralist" error is the selection of a button that is correct for a stimulus other than the one presented, and "Task I intrusions" involve selections of buttons used in Task I. Task I intrusions involving response to a stimulus in Task II with the button that was correct in Task I were designated "Old habit errors" and are shown separately for the A-C conditions in Table 5.

The major trends in Table 5 may be summarized briefly. The frequencies of omissions, intralist errors, and Task I intrusions on the D-C pairs tend to decrease with increasing amounts of first-task practice. No consistent trends for these three types of error are apparent for the A-C pairs. The differences between the D-C and A-C totals for both omissions and intralist errors tend to increase with increasing first-task practice. The differences between D-C and A-C totals for Task I intrusions tend to decrease, if anything, with increasing first-task practice. No marked trend for the old habit errors is apparent. From these data, considered by themselves, it appears that either omissions or intralist errors constitute the most sensitive index of associative interference and that Task I intrusions are either inversely related or unrelated to associative interference.

Experiment III

The failure in Exp. II to show a decrease in the A-C correct anticipation means, with increase in first-task practice, is presumably due to the effect on A-C performance of such nonspecific factors as warm-up and learning-to-learn. An experiment that varied amount of first-task practice within *Ss*, by giving different amounts of practice on subsets of first-task pairs and then requiring *Ss* to learn a new response to each of these stimuli, should show steadily increasing interference as a function of amount of first-task practice. Such a design would assure that *Ss* have had equal amounts of total practice with the task for each of the different amounts of first-task practice. Experiment III is an attempt to demonstrate this hypothesis.

TABLE 6

Light-Button Pairings for Task I and Task II (Exp. III and Exp. IV)

Pair	Stimuli	Correct Buttons	
		Task I	Task II
A	Red	4	1
	White	7	3
B	Blue	2	5
	Yellow	10	6
C	Green	12	9
	Violet	8	11

Method

Apparatus.—The apparatus for Exp. III was the same as that for Exp. I and II except that the response unit consisted of 12 push-button switches instead of 9.

Subjects, design, and procedure.—The *Ss* were 30 children from Grade 4 of the Ottumwa Public School System.[4] All *Ss* were given six stimuli in Task I and were required to associate a different one of 6 of the 12 buttons with each stimulus. In Task II, they were given these same six stimuli and were required to associate a different one of the other six buttons with each stimulus. The light-button pairs for the two tasks are shown in Table 6. The amount of Task I practice was varied within *Ss* by presenting, in Task I, two stimuli 5 times, two others 10 times, and a third pair 20 times. Six subgroups of five *Ss* each were randomly formed in order to counterbalance for the intrinsic difficulty of the S-R pairs. All six ways of combining the three stimulus pairs (A, B, and C of Table 6) with the three levels of first-task training were thus represented.

The procedures for Exp. III were the same as those for Exp. I and II with the following exceptions. During Task I, a total of 70 stimulus presentations was administered. In each successive block of 14 presentations, the first pair of stimuli was presented once; the second pair, twice; and the third pair, four times. Within a block of 14 presentations, the order of presenting the stimuli was random with the restriction that no stimulus appear twice or more successively. For each of the 5 blocks of 14 presentations, a different order was used. During the interpresentation interval following the last stimulus presentation of Task I, *E* unobtrusively turned a selector switch to re-pair the lights and the buttons for Task II. Thus, *S* was started on Task II without warning. Twelve trials on Task II were administered, where each trial included one presentation of each of the six stimuli.

[4] See Footnote 2.

TABLE 7

MEANS AND SD's OF NUMBER OF CORRECT RESPONSES IN TASK I FOR DIFFERENT CONDITIONS OF TRAINING

Conditions of Training	Experiment III				Experiment IV			
	First 5 Presentations		Last 5 Presentations		First 5 Presentations		Last 5 Presentations	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
5*	3.23	2.32	3.23	2.32	3.93	2.08	3.93	2.08
10	2.80	2.58	6.13	2.86	2.67	2.06	6.20	3.32
20	2.65	1.79	8.27	2.65	2.93	2.00	8.37	1.78

* For Cond. 5, the first five presentations were the same as the last five.

Results

Task I.—The first two columns of Table 7 give the means and *SD*'s of the number of correct responses to the first five presentations for each of the three conditions of first-task training. It is apparent that there are fewer correct responses to the first five presentations for the stimuli presented 20 times in Task I than for those presented five times. An analysis of variance of these scores yielded an *F* ratio significant at only the 10% level. The meaning of these differences will be discussed in connection with the results of Exp. IV.

The third and fourth columns of Table 7 present the means and *SD*'s of the number of correct responses to the last five presentations for the three conditions of first-task practice. The final level of performance is seen to be an increasing function of the number of previous presentations. Analysis of variance, with amount of first-task practice and stimulus pairs as within-*S* factors (**3**, Type II, p. 273), indicates that these differences are highly significant ($P < .001$).

Task II.—The first three columns of Table 8 give the means for Task II of the different types of error that were considered in Exp. II. As expected, the number of total errors increases with increasing number of first-task presentations. Analysis of variance (**3**, Type II, p. 273), however, showed the difference in these means to be

TABLE 8

MEANS OF DIFFERENT TYPES OF ERROR IN TASK II FOR DIFFERENT CONDITIONS OF TRAINING

Type of Error	Experiment III			Experiment IV		
	Task I Presentations			Task I Presentations		
	5	10	20	5	10	20
Omissions	4.27	4.97	5.63	5.17	5.47	7.97
Intralist errors	5.07	5.43	5.17	5.03	6.10	4.60
Task I intrusions	3.27	2.60	3.00	5.20	4.57	5.07
Old habit errors	(0.93)	(0.60)	(0.80)	(1.43)	(0.97)	(1.20)
Total errors	12.60	13.00	13.80	15.40	16.13	17.63

nonsignificant ($P > .20$). Examination of the means for response omissions indicates that they increase with an increase in the number of Task I presentations. This trend was found to be significant at only the 10% level. Intralist errors, Task I intrusions, and old habit errors do not manifest a systematic change with variation in degree of prior learning, and in each case, the differences among means were nonsignificant ($P > .20$).

Experiment IV

On the assumption that omissions best reflect interference in the A-B, A-C design, Exp. IV was conducted in an attempt to increase the number of omissions by reducing the length of the anticipation interval during Task II. Experiment I was replicated with another group of 30 fourth-grade children from the same school,[5] the sole procedural change involving a reduction in the anticipation interval in Task II from 2 sec. to 1½ sec.

Results

Task I.—The means and *SD*'s of the correct responses to the first and to the last five presentations are given in the last four columns of Table 7. The *Ss* were again found to perform better to the first five presentations of the stimuli that were presented only five times than they did to the stimuli presented either ten or twenty times. Analysis of variance showed the differences among the three means to be significant ($P < .01$). The number of correct responses to the last five presentations again was found to increase significantly ($P < .001$) with an increase in the number of previous presentations.

Task II.—The means of the different types of error in Task II are shown in the last three columns of Table 8. Total errors were again found to increase with increase in number of Task I presentations, and

[5] See Footnote 2.

TABLE 9

Analysis of Variance of Total Errors in Task II for Exp. IV

Source	*df*	*MS*	*F*
Between *Ss*	29		
T × P (b)	2	10.35	<1.00
Error (b)	27	41.83	
Within *Ss*	60		
Training conditions (T)	2	38.88	3.82*
Stimulus pairs (P)	2	64.68	6.35**
T × P (w)	2	2.32	<1.00
Error (w)	54	10.19	

* $P < .05$.
** $P < .005$.

omissions again appear to reflect the increasing interference somewhat more sensitively than does any one of the other error measures.

The analysis of variance conducted on number of total errors is summarized in Table 9. Stimulus pairs were found to differ significantly ($P < .005$) reflecting the greater ease in learning Pair A (Table 1) than in learning the other two. The means for the three conditions of training differed at the 5% level of confidence. Using the error term from this analysis, a t test was conducted for the means of the 5 and 20 conditions. The resultant t was 2.71, $P < .01$.

The same type of analysis was conducted over each of the other four types of error. No significant differences among means were found for old habit errors, intralist errors, and Task I intrusions ($P > .10$). The analysis of omissions, however, yielded a significant F ratio for the three conditions ($P < .001$), and t tests showed that Cond. 5 and 10 each differed significantly from Cond. 20 ($P < .001$), but did not differ significantly from each other ($P > .20$).

Discussion

The results of Exp. III and IV are consistent in showing amount of interference

in Task II to be an increasing function of amount of first-task practice, when nonspecific transfer is experimentally controlled. In Exp. III, the results were not statistically reliable. In Exp. IV, the results were in the same direction, and, apparently due to the reduction in anticipation time, were more reliable.

The finding that Task I performance on the first five presentations differed for the three conditions of practice requires some consideration. The five presentations of the stimuli in Cond. 5 were spread throughout the 70 stimulus presentations of Task I, while the first five presentations of the stimuli in Cond. 20 were given in the first 18 presentations. Thus, in Cond. 5, *S*s were receiving some of the presentations late in Task I, while in Cond. 20, they were receiving all the presentations early in the task. The finding, then, represents an instance of positive transfer within a single task, the positive transfer presumably due to amount of prior practice on the task. The fact that it is potent enough to be statistically significant in such a relatively crude test emphasizes the need to control such effects if one wishes to assess interference due to the prior learning of incompatible associations.

It is of interest that omissions in Task II appear to reflect the interference of Task I training more directly than any other type of error measure considered. The relative infrequency of Task I intrusions is consistent with the results of verbal paired-associate experiments. In the A-B, A-C paradigm, *S*s apparently inhibit the old responses rather easily, but this inhibition is not always followed by the immediate selection of a new response. It frequently results in a temporary blocking of any overt response.

If this analysis is correct, one should expect the reaction time or latency, even of a correct response, to increase with increasing strength of incorrect competing responses. In Exp. III and IV, for example, if the latencies of all responses in Task II had been obtained, longer latencies of correct responses would be expected to those stimuli presented 20 times in Task I than to those presented only five times.

Summary

Four experiments were conducted to study interference in the A-B, A-C design with motor paired-associate tasks. In Exp. I and II, an intra-*S* experimental design, with an A-B, D-C control condition, was devised in order to provide statistical control of warm-up and learning-to-learn. In Exp. I, 18 grammar school *S*s were given preliminary training on a motor paired-associate task consisting of three S-R pairs. In the second (criterial) task, *S*s were required to learn three new responses to these same stimuli, and three new responses to three additional stimuli. The results, in terms of number of correct responses in each condition, showed significant interference for the A-C condition. In Exp. II, 18 grammar school children were assigned to each of three experimental groups differing only in amount of Task I practice. It was found that the amount of interference on the A-C, relative to the D-C, condition was an increasing function of amount of Task I practice.

Experimental control of nonspecific transfer was obtained in Exp. III and IV. In Exp. III, 30 grammar school children were required to learn to push the appropriate one of 12 available response buttons to each of six different stimuli. In Task I, they were given 5 presentations of two of the S-R pairs, 10 of another two, and 20 of the third two. In Task II, they were required to learn to push the appropriate one of the six buttons not used in Task I to each of the same stimuli. Number of errors in Task II was found to increase with number of Task I presentations, although the differences were nonsignificant. An additional 30 *S*s were used in Exp. IV which differed from Exp. III only in that a shorter anticipation interval was used. The findings of Exp. III were replicated in Exp. IV at a statistically significant level of confidence.

References

1. Atwater, S. K. Proactive inhibition and associative facilitation as affected by degree of prior learning. *J. exp. Psychol.*, 1953, **46**, 400–404.
2. Bruce, R. W. Conditions of transfer of training. *J. exp. Psychol.*, 1933, **16**, 343–361.
3. Gibson, E. J. Retroactive inhibition as a function of degree of generalization between tasks. *J. exp. Psychol.*, 1941, **28**, 93–115.

4. Greenberg, R., & Underwood, B. J. Retention as a function of stage of practice. *J. exp. Psychol.*, 1950, **40**, 452–457.
5. Hamilton, C. E. Length of interval between two tasks and performance on the second task. *J. exp. Psychol.*, 1950, **40**, 613–621.
6. Lindquist, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1953.
7. McGeoch, J. A., & Underwood, B. J. Tests of the two-factor theory of retroactive inhibition. *J. exp. Psychol.*, 1943, **32**, 1–16.
8. Morgan, R. L., & Underwood, B. J. Proactive inhibition as a function of response similarity. *J. exp. Psychol.*, 1950, **40**, 592–603.
9. Thune, L. E. The effect of different types of preliminary activities on subsequent learning of paired-associate material. *J. exp. Psychol.*, 1950, **40**, 423–438.
10. Underwood, B. J. Retroactive and proactive inhibition after five and forty-eight hours. *J. exp. Psychol.*, 1948, **38**, 29–38.
11. Underwood, B. J. Spontaneous recovery of verbal associations. *J. exp. Psychol.*, 1948, **38**, 429–439.
12. Underwood, B. J. Proactive inhibition as a function of time and degree of prior learning. *J. exp. Psychol.*, 1949, **39**, 24–34.
13. Underwood, B. J. Associative transfer in verbal learning as a function of response similarity and degree of first-list learning. *J. exp. Psychol.*, 1951, **42**, 44–53.
14. Underwood, B. J. Studies of distributed practice: III. The influence of stage of practice in serial learning. *J. exp. Psychol.*, 1951, **42**, 291–295.
15. Ward, L. B. Reminiscence and rote learning. *Psychol. Monogr.*, 1937, **49**, No. 4 (Whole No. 220).

(Received September 13, 1957)

Journal of Experimental Psychology
Vol. 56, No. 2, 1958

RECOGNITION OF NOVEL VISUAL CONFIGURATIONS WITH AND WITHOUT EYE MOVEMENTS[1]

CRAIG M. MOONEY

Defence Research Medical Laboratories, Toronto

Earlier studies of perceptual closure (**4, 5, 6**) have shown that the provision of ample viewing time with opportunity for multiple visual fixations does not facilitate the recognition of incomplete representations of familiar objects; such pictures are as effectively perceived with a single fixation, whether that fixation be brief or prolonged.

The finding is surprising in view of the depleted and enigmatic nature of the visual representations employed; one might have supposed that these called for visuo-neural redintegrative processes which are, according to theorists like Hebb (**2**), Lashley (**3**), and Gibson (**1**), essentially dependent on prosecutory eye movements.

It is more difficult to test the contention—central in Hebb's thinking, for example—that eye movements are essential for that perceptual learning whereby the unfamiliar object becomes familiar enough to be perceived at a glance. Nevertheless, in view of the foregoing finding with enigmatic closure presentations, it is appropriate to raise the question: Are multiple visual fixations essential, or significantly superior to a single brief fixation, in learning to recognize an unfamiliar object?

In any experimental studies addressed to such a question, it will be necessary to specify the perceptual criterion—whether it shall be, for example, denomination of the object as a member of a class, or its identification as a particular one of a set of objects, or its delineation or description by recall, or its recognition as a particular object previously seen.

As a simple beginning, and most directly to realize the intention of pitting the single glance against visual scanning, the last-mentioned method of recognition has been employed in the present study. Novel visual configurations have been displayed in ways designed to reveal whether their recognition on a second occasion of experience is significantly enhanced when their initial observation, on the first occasion of experience with them, entails eye movements rather than a single brief fixation.

METHOD

Two kinds of test items were used, as shown in Fig. 1. These were contrived pragmatically to be amorphous (Type I), and structured (Type II). Both were produced in photographically positive and negative states on 35-mm. slides.

The *S* sat to the right of a projecting apparatus in a dark room viewing displays 2 ft. wide on a screen 6 ft. away. The *S* was instructed to look "hard" at each configuration as it was displayed in order to be able to recognize it later. When the requisite items had each been shown once, *S* was told that they would be mixed in with a number of other items of the same kind which he had not seen before, and that he would be required to identify each one as it was shown by saying "Yes" or "No"; that by "Yes" he would mean that it was one of the items he had originally seen, and by "No" that it was a new one he had not previously seen. "Yes" and "No" scores were then separately tallied.

Four experiments were performed, utilizing experimentally naive adult soldiers as *S*s in each. The basic procedure was to test *S*s with different lots of items presented, respectively: (*a*) in the same spatial orientation and graphic (black-white) state in which they had first been shown; (*b*) in reversed orientation (left to right, or top to bottom); (*c*) in reversed state (blacks and whites

[1] DRML Report No. 142-1, Project No. D77-94-35-29 (HR No. 160).

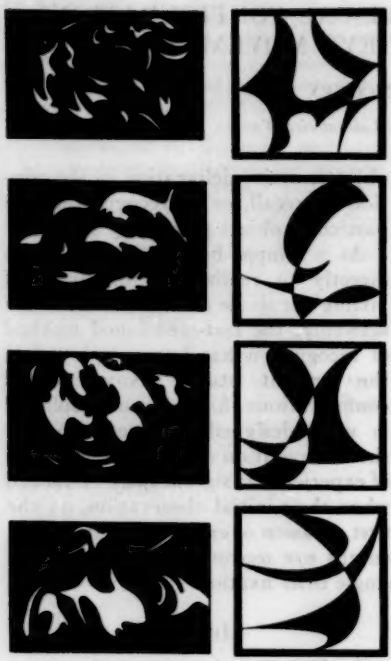

FIG. 1. Examples of Type I (amorphous) and Type II (structured) items.

reversed); (*d*) in reversed orientation and state. Two contrasting exposure speeds were employed with each method of presentation: usually .04 sec. and 10 sec. (In practice, achieving a left-to-right reversal in orientation simply required turning the upright slide around sideways from front to back in the projector; a top-to-bottom reversal in orientation, turning it over from front to back; a reversal in state, substituting a negative slide for a positive one.) In Exp. I, Cond. *a* and *b* were utilized; in Exp. II, Cond. *c* and *d*; and in Exp. III one set of *S*s was tested under Cond. *a*, *b*, *c*, and *d*. Experiment IV was not exactly analogous to these, but dealt with related questions.

EXPERIMENT I

The purpose was to ascertain whether the recognition of novel configurations on a second occasion of experience with them depends upon the amount of viewing time afforded on the first occasion, under two test conditions: where they reappear in the same orientation as initially shown, and where they reappear in reversed (left-right) orientation.

Method.—Three groups of 8 *S*s were used. Each group was tested with two lots of items. Initial presentation of items and subsequent testing with them employed the same exposures; these were, for the respective *S*-groups: .04 sec., 1 sec., and 10 sec. Items in one lot were presented for test in the same orientation in which they had first been shown; items in the second lot were presented for test in reversed (left-right) orientation to that in which they had first been shown. A lot consisted of 10 randomly assigned Type I (key) items mixed with 20 similar ones. A factorial design was employed to balance exposures, item lots, and sequence.

Each *S* was shown the 10 (key) items one after the other, for the designated exposure, at about 5-sec. intervals, and was tested immediately thereafter by being shown the 30 items, at the exposure and rate initially employed. The *S*'s scores, for each lot of items, were the number of correct "Yes" and "No" answers.

Results.—The findings are summarized in Table 1. Separate analyses of variance for the two categories of answers revealed no significant differences attributable to the different exposures used, nor to the left-right reversal in orientation.

EXPERIMENT II

The purpose was to ascertain whether the recognition of novel configurations on a second occasion of experience with them depends upon the amount of viewing time afforded on the first occasion, under two test conditions: where they reappear in their negative (black white) state in the same orientation as initially shown and where they reappear in their negative state reversed in (left-right) orientation.

Method.—Three groups of 8 *S*s were used. Items, exposures, and procedures were identical with those in Exp. I except that *S*s were tested with all items in both lots presented as negative, that is, in reversed (black-white) state.

TABLE 1

RECOGNITION OF NOVEL CONFIGURATIONS UNCHANGED IN (BLACK–WHITE) STATE PRESENTED IN ORIGINAL AND REVERSED (LEFT–RIGHT) ORIENTATION AT DIFFERENT EXPOSURES

Method of Presentation	Exposure (Sec.)		*Ss*	Items		Correct		Per Cent	
	Initial	Test		Key	New	Key	New	Key	New
Original orientation; original state	.04	.04	8	10	20	54	145	67	91
	1	1	8	10	20	56	148	70	92
	10	10	8	10	20	60	138	75	86
			24	10	20	170	431	71	90
Reversed orientation; original state	.04	.04	8	10	20	53	143	66	89
	1	1	8	10	20	50	141	62	88
	10	10	8	10	20	46	140	57	87
			24	10	20	149	424	62	88

Results.—The findings are summarized in Table 2. The analyses of variance revealed that significant differences cannot be attributed to any of the factors. The differences associated with reversal in orientation and state are such as could occur by chance.

EXPERIMENT III

The purpose was to ascertain whether the recognition of novel configurations on a second occasion of experience with them depends upon the amount of viewing time afforded on the first occasion, under four test conditions: where they reappear in original and reversed (black-white) state, in, respectively, original and reversed (left-right) orientation.

Method.—Three groups of 16 *Ss* were used. With each group a different exposure was used. These exposures were, referring to initial and subsequent presentations, respectively: .04 sec. and .04 sec.; 5 sec. and .04 sec.; 5 sec. and 5 sec. Each *S* was tested with four lots of items, as

TABLE 2

RECOGNITION OF NOVEL CONFIGURATIONS REVERSED IN (BLACK–WHITE) STATE PRESENTED IN ORIGINAL AND REVERSED (LEFT–RIGHT) ORIENTATION AT DIFFERENT EXPOSURES

Method of Presentation	Exposure (Sec.)		*Ss*	Items		Correct		Per Cent	
	Initial	Test		Key	New	Key	New	Key	New
Original orientation; reversed state	.04	.04	8	10	20	45	126	56	79
	1	1	8	10	20	44	130	55	81
	10	10	8	10	20	39	147	49	92
			24	10	20	128	403	53	84
Reversed orientation; reversed state	.04	.04	8	10	20	45	121	56	76
	1	1	8	10	20	37	137	46	86
	10	10	8	10	20	23	150	29	94
			24	10	20	105	408	44	85

TABLE 3

RECOGNITION OF NOVEL CONFIGURATIONS PRESENTED IN ORIGINAL AND REVERSED (BLACK-WHITE) STATE AND (LEFT-RIGHT) ORIENTATION AT DIFFERENT EXPOSURES

Method of Presentation	Exposure (Sec.)		*Ss*	Items		Correct		Per Cent	
	Initial	Test		Key	New	Key	New	Key	New
Original orientation; original state	.04	.04	16	4	8	38	107	59	84
	5	.04	16	4	8	46	104	72	81
	5	5	16	4	8	50	114	78	89
			48	4	8	134	325	70	85
Reversed orientation; original state	.04	.04	16	4	8	32	112	50	87
	5	.04	16	4	8	34	100	53	78
	5	5	16	4	8	35	108	55	84
			48	4	8	101	320	53	83
Original orientation; reversed state	.04	.04	16	4	8	42	101	66	79
	5	.04	16	4	8	35	109	55	85
	5	5	16	4	8	32	106	50	83
			48	4	8	109	316	57	82
Reversed orientation; reversed state	.04	.04	16	4	8	34	101	53	79
	5	.04	16	4	8	25	106	39	83
	5	5	16	4	8	26	102	41	80
			48	4	8	85	309	44	80

follows: (*a*) with the key items in their original orientation and state; (*b*) with the items in reversed (left-right) orientation; (*c*) with the items in reversed (black-white) state; (*d*) with the items in reversed orientation and state. An item lot consisted of four randomly assigned Type II key items mixed with eight similar ones. Other procedures were the same as in the preceding experiments.

Results.—The findings are summarized in Table 3. The analyses of variance revealed that differences attributable to the different exposures are not significant. Apart from this, performance is significantly depressed by (left-right) reversal in orientation, and by change from positive to negative state, and these factors are independent of each other. It will be noted, by a comparison of Tables 1 and 2 with the first and second halves of Table 3, that the findings in Exp. I and II are not at variance with the corresponding findings in Exp. III.

EXPERIMENT IV

The purpose was to ascertain whether the recognition of novel configurations on a second occasion of experience with them depends upon repeated single glances or continuous viewing time afforded on the first occasion, using structured and amorphous configurations, under two test conditions: where they reappear in their original orientation, and where they reappear upside down.

Method.—Two groups of 20 *Ss* were used. Each was tested with two lots of items—one lot being Type I, the other Type II—with circular frames replacing the rectangular ones used in the earlier experiments. This entailed initially showing each item in one lot for 10 sec., and showing each item in the other lot 10 successive

TABLE 4

RECOGNITION OF AMORPHOUS AND STRUCTURED NOVEL CONFIGURATIONS PRESENTED IN ORIGINAL AND UPSIDE-DOWN ORIENTATIONS AT DIFFERENT EXPOSURES

Method of Presentation	Type of Configuration	Initial Exposure*	*Ss*	Items		Correct		Per Cent	
				Key	New	Key	New	Key	New
Original orientation	Structured	10	10	4	8	28	71	70	89
	Amorphous	10	10	4	8	25	74	62	92
	Structured	10 × .1	10	4	8	30	75	75	94
	Amorphous	10 × .1	10	4	8	27	72	67	84
			20	8	16	110	292	69	91
Upside down	Structured	1 × 10	10	4	8	21	65	52	81
	Amorphous	1 × 10	10	4	8	17	74	42	92
	Structured	10 × .1	10	4	8	24	78	60	97
	Amorphous	10 × .1	10	4	8	17	67	42	84
			20	8	16	79	284	49	89

* All test exposures were .04 sec.

times (at exposures of .1 sec., at 1 sec. intervals) and subsequently testing by exposing items for .04 sec., in both lots. One group of *Ss* was tested with both lots of items presented in their original orientation; the other group was tested with both lots of items presented upside down. Other procedures were the same as in the preceding experiments.

Results.—The findings are set out in Table 4. Significant differences cannot be attributed to the two methods of initial presentation employed, nor to the two types of configurations used. There is a significant difference between the original and the upside-down presentation.

DISCUSSION

The gist of the findings is that the subsequent recognizability of novel visual configurations immediately following the first occasion of experience with them is no greater when that occasion affords viewing time with opportunity for scanning eye movements than when it permits but a single glance.

Recognition of configurations seen previously was about 70% effective when they appeared in their original orientation and state; and nonrecognition of interspersed configurations not previously shown was about 85% effective. Such levels of accomplishment are consistent with those found in earlier studies —Achilles; Lehmann; Seward; and Burtt and Dobell—as cited by Woodworth (7). The initial provision of lengthy viewing periods (1, 5, 10 sec.) or of a succession of brief glances (10 exposures of .1 sec. at 1-sec. intervals) did not improve the level of subsequent recognition over that achieved with an initial glance of .04 sec.

Performance declined significantly when the configurations were presented for recognition in reversed orientation and state. For example, *of all configurations correctly recognized,* 30% occurred under the condition where they were displayed in their original orientation and state; 25% under the condition where they were reversed in either of these respects; and 20% where they were reversed in both. In this latter condition, and also in the instance where configurations were turned upside down, performance was at the chance level. These declinations were not related to the different modes of viewing initially provided.

Subsumed in the experiments, through provision of amorphous and structured

configurations and their presentation in changed orientations and states, was the idea that, if the scanning method of observation was doing any kind of mnemonic visual work by utilizing the elements and positions of the configurations (a labor presumably not so feasible in a single brief glance), this might be reflected in the results.

The want of any such evidence, and the conclusion that the memories (as defined in these experiments) for novel configurations, consequent on the first occasion of experience with them, are effectually laid down with a single brief glance, suggests that the requisite learning is not of a kind that can be prosecuted with scanning eye movements. This may be because the configurations, by virtue of individual novelty and uniqueness, cannot be made familiar through associative, comparative processes of visual inspection and cognitive annotation, but are notable only by virtue of their idiomatic "thing-ness," for which the sufficient and, apparently, only possible "trace" is achieved with the single brief glance.

Summary

Four experiments were addressed to the question of the superiority of scanning eye movements over a single brief glance in learning to recognize novel visual configurations. The criterion of perceptual learning was the recognition of configurations on a second occasion of experience as ones that had been observed on the first occasion. Testing entailed showing "key" items intermixed with new items.

The experiments were designed to permit comparison between recognitions achieved with initial exposures of several seconds' duration and those achieved with a single brief glance. Both amorphous and structured configurations were employed; and the key configurations were presented for test in both original and reversed (left-right) orientations and original and reversed (black-white) graphic states.

The findings revealed that no advantage accrued to visual inspection (viewing time with opportunity for multiple visual fixations) over the single brief glance (a single fixation of a fraction of a second) under any of the various experimental conditions. The findings suggest that the kind of perceptual learning entailed is not of a kind that can be prosecuted with eye movements, due presumably to the novel and idiomatic nature of the configurations.

References

1. Gibson, J. J. *The perception of the visual world.* Boston: Houghton Mifflin, 1950.
2. Hebb, D. O. *The organization of behaviour.* New York: Wiley, 1949.
3. Lashley, K. S. Dynamic processes in perception. In *Brain mechanisms and consciousness: A symposium.* Springfield: Thomas, 1954. Pp. 422–443.
4. Mooney, C. M. Closure with negative afterimages under flickering light. *Canad. J. Psychol.*, 1956, **10**, 191–199.
5. Mooney, C. M. Closure as affected by viewing time and multiple visual fixations. *Canad. J. Psychol.*, 1957, **11**, 21–28.
6. Mooney, C. M. Closure as affected by configural clarity and contextual consistency. *Canad. J. Psychol.*, 1957, **11**, 80–88.
7. Woodworth, R. S. *Experimental psychology.* New York: Holt, 1938.

(Received September 20, 1957)

Journal of Experimental Psychology
Vol. 56, No. 2, 1958

VERBAL CONDITIONING AND EXTINCTION WITH VERBAL AND NONVERBAL REINFORCERS [1]

ARNOLD H. BUSS
University of Pittsburgh

IRMA R. GERJUOY
Carter Hospital

AND JACK ZUSMAN
Albert Einstein Medical School

A number of recent studies (**1, 2, 3, 4, 6**) have employed a modified operant procedure in the conditioning of verbs of the past tense. The *S* is shown a card with six pronouns and a verb and is instructed to make up a sentence using the verb and one of the pronouns. After a free responding period *S* is reinforced for using first-person pronouns in his sentences. It has been found that *E*'s saying "Good" is an effective reinforcer—i.e., the number of sentences starting with "I" or "We" increases with trials.

The experiment most relevant to the present investigation is the one by Cohen, Kalish, Thurston, and Cohen (1). They had an 80-trial acquisition period with "Good" as the reinforcer, followed by 80 trials of extinction during which *E* said nothing. It was found that the frequency of first-person pronoun responses rose steadily throughout acquisition; however, during extinction the frequency of first-person pronouns did not drop, but rather remained at about the same level as at the end of acquisition.

Cohen et al. noted that secondary reinforcement offers a possible explanation for the absence of extinction. During extinction *E* continued to record and present successive cards, and these events might have carried enough reinforcement properties to maintain the previously acquired response. This hypothesis might be tested by preventing *S* from seeing *E* and having the cards presented by an apparatus. However, to the writers' knowledge such a study has not been conducted.

There is an alternative explanation to the secondary reinforcement hypothesis. It should be noted that the reinforcer "Good" has no analogue in the infrahuman studies from which virtually all of the data on instrumental conditioning and extinction have been obtained. It is possible that with human *S*s a verbal reinforcer during acquisition leads to an extremely slow decrement in response during extinction. This hypothesis might be tested by comparing extinction after "Good" with extinction after a reinforcer that is more analogous to those used in infrahuman studies. The present study attempts to make this test, using the nonverbal reinforcers cigarettes and candy in opposition to the verbal reinforcer "Good."

METHOD

Design.—All *S*s were instructed to make up one sentence for each of 160 cards. Each sentence constituted a trial. In order to ascertain the operant level of responding with the pronouns I and We, the *S*s received no reinforcement for the first 20 trials. The *E* recorded each pronoun choice for these and all subsequent trials.

The acquisition period consisted of Trials 21–80. For Group I starting with Trial 21 the *E* said "Good" after every sentence containing I or We. For Group II *E* gave a poker chip after every sentence starting with I or We beginning with Trial 21. The chips were to

[1] This research was conducted while the writers were on the staff of Carter Memorial Hospital, Indianapolis. The writers acknowledge the considerable help of Herbert Gerjuoy.

be turned in for cigarettes or candy. Chips were used both as a matter of convenience and in order to decrease any satiation effects. It is possible that chips alone would be reinforcing, cigarettes and candy being superfluous. Therefore a third group was added. For Group III *E* gave a poker chip for every sentence starting with I or We beginning with Trial 21. These chips were not to be exchanged for cigarettes or candy; they were merely returned to *E* at the end of the session.

The extinction period, Trials 81–160, was the same for *S*s in all groups. No reinforcement was given for any response during that period.

Materials.—The stimulus materials were 160 3 × 5-in. white index cards. In the center of each card was a past-tense verb, and beneath each verb were six pronouns: I, We, He, She, You, They. The order of pronouns was varied randomly from card to card, and a different verb appeared on each card.

Subjects.—The *S*s were 78 students from introductory psychology classes at Indiana University and 78 patients in a neuropsychiatric hospital. There were equal numbers of male and female *S*s in both student and patient groups. The requirements for serving as *S* were reading competence and ability to follow the instructions.

The 156 *S*s were randomly assigned to three treatment groups of 52 *S*s each, with the restriction of equal numbers of male patients, female patients, male students and female students in each group.

Procedure.—The initial instructions for all *S*s were as follows: "This is a study of how people make up sentences. I will show you a bunch of cards one at a time. Each card will have a word in the center and six words on a line underneath. What I want you to do is to make up one sentence for each card, using the word in the middle and one of the words on the line below. Try this one [*S* was given the sample card]. That's fine. That's what you will be doing throughout the whole experiment. Your sentences may be long or short, it doesn't matter. You will not be timed, but there are a lot of cards to go through, so go as fast as you can. Do you have any questions?"

The *E* answered questions by repeating the appropriate part of the instructions. The *S* was then given the stack of cards to flip himself. The *E* used a code to record the *S*'s responses.

The *S*s in Group I received no further instructions. The *E* said "Good" in a monotone during Trials 21–80 after each sentence starting with I or We. If *S* questioned *E*'s response, he was told that no further questions would be answered.

After Trial 20 in Group II *S*s were instructed: "Now sometimes you will get a chip after a sentence and sometimes you will not. When we are all through you may turn in your chips for cigarettes or candy, whichever you prefer. You may have one cigarette or one piece of candy for every three chips. In other words, the chips represent cigarettes and candy which I will give you as your reward for taking part in the experiment."

The cigarettes and candy were in plain view of *S*s in this group. The *E* gave each *S* in this group one chip after each sentence in which he used I or We in Trials 21–80.

The *S*s in Group III were given the following instructions after Trial 20: "Now sometimes you will get a chip after a sentence and sometimes you will not." The *E* gave each *S* one chip after each I or We response in Trials 21–80. Neither cigarettes nor candy were in sight.

After Trial 160 all *S*s were questioned about the nature of the experiment. The *E* asked each *S* if he knew what the experiment was about. If he could answer correctly, he was excused. Otherwise he was questioned about whether he knew what the connection was between what he said and what *E* did (say "Good" or give chip).

Results

Acquisition.—The acquisition data are plotted in Fig. 1. The initial block of 20 trials represents the free responding period. The differences among groups during the free responding period appear to be small and random, and the absence of system-

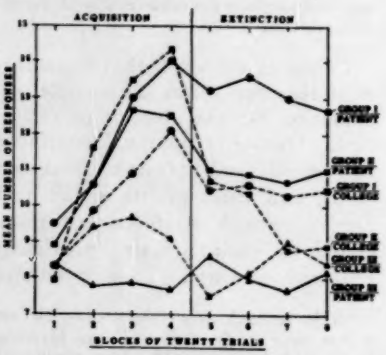

FIG. 1. Acquisition and extinction of first-person pronoun responses.

TABLE 1

ANALYSIS OF VARIANCE OF ACQUISITION DATA

Source	*df*	*MS*	*F*
Groups (G)	2	526.76	8.98**
Sex (S)	1	24.26	.41
Populations (P)	1	2.20	.04
G × S	2	180.07	3.07*
G × P	2	102.89	1.75
S × P	1	.35	.01
G × S × P	2	31.00	.53
Error (b)	144	58.69	
Trial Blocks (T)	3	299.99	31.88**
T × G	6	77.10	8.19**
T × S	3	10.31	1.10
T × P	3	14.05	1.49
T × G × S	6	9.49	1.01
T × G × P	6	9.18	.98
T × S × P	3	1.24	.13
T × G × S × P	6	8.03	.85
Error (w)	432	9.41	
Total	623		

* $P < .05$.
** $P < .001$.

atic effects was confirmed by an analysis of variance.

The six acquisition curves in Fig. 1 vary considerably in shape. The curves of both patients and college students in Group I ("Good") rise steadily throughout acquisition. Similarly, the curves of both patients and college students in Group II (Cigarettes-Candy) have clearly positive slopes. On the other hand the curves of both patients and college students in Group III (Chips alone) are relatively flat, indicating an absence of conditioning.

An analysis of variance (**5,** modified Type III) of the acquisition data is shown in Table 1. The significant Blocks × Groups effect confirms the trend seen in Fig. 1: the curves are significantly different in shape. The significant Groups effect merely adds that the curves also differ significantly in height. The significant Groups × Sex effect indicates that the difference between men and women in total number of responses varied from group to group.

Extinction.—The extinction data are plotted in the right half of Fig. 1. The acquisition curves of Group III indicate that virtually no learning had occurred. Therefore the "extinction" data of Group III could have little meaning, and they do not enter into the following analysis.

Of the four remaining curves, one stands out. This is the Group II college curve, whose slope appears to be somewhat steeper than the slopes of the other three curves. In order to test differences among the extinction curves, an analysis of variance was performed on the extinction data plus the last block of acquisition trials. The results are presented in Table 2. The *F*'s in the upper half of the table involve the height of the curves, and those in the bottom half involve the shape of the curves.

Inspection of Table 2 reveals that the curves do not differ significantly

TABLE 2

ANALYSIS OF VARIANCE OF EXTINCTION DATA

Source	*df*	*MS*	*F*
Groups (G)	1	160.62	1.56
Sex (S)	1	27.23	.26
Populations (P)	1	312.23	3.03
G × S	1	5.82	.06
G × P	1	121.16	1.17
S × P	1	4.63	.04
G × S × P	1	35.03	.34
Error (b)	96	103.18	
Trial Blocks (T)	4	108.56	14.46**
T × G	4	20.46	2.72*
T × S	4	15.13	2.01
T × P	4	19.84	2.64*
T × G × S	4	3.09	.41
T × G × P	4	12.95	1.72
T × S × P	4	5.72	.76
T × G × S × P	4	4.96	.66
Error (w)	384	7.51	
Total	519		

* $P < .05$.
** $P < .001$.

TABLE 3

ANALYSIS OF COVARIANCE OF EXTINCTION SLOPES

Source	*df*	*MS*	*F*
Groups (G)	1	2.44	2.39
Sex (S)	1	2.45	2.40
Populations (P)	1	5.05	4.95*
G × S	1	.55	.54
G × P	1	2.01	1.97
S × P	1	.27	.26
G × S × P	1	.38	.37
Error (w)	95	1.02	
Total	102		

* $P < .05$.

in height but there are significant differences in shape. The significant Blocks × Groups interaction indicates that the combined Group I curves are significantly flatter than the combined Group II curves. Since this comparison is not represented in Fig. 1, the patients' and college students' data within each group have been combined. The means for the last five blocks of trials are as follows: Group I—13.0, 11.8, 12.1, 11.6, 11.6; and Group II—13.6, 10.8, 10.6, 9.8, 9.9. These data demonstrate that the "Good" group extinguished slower than the Cigarettes-Candy group.

The significant Blocks × Populations interaction involves combining patients' data across groups and comparing them with the combined college students' data. The means for the last five blocks of trials are: patients—13.3, 12.1, 12.2, 11.8, 11.8, and college students—13.2, 10.6, 10.5, 9.6, 9.7. Thus the patients extinguished significantly slower than the college students.

The Blocks × Groups × Populations interaction is not significant, which suggests that (*a*) Fig. 1 is deceptive and the extent to which the Group II college curve is different from the other three curves is merely random, or (*b*) extraneous variance is obscuring the underlying relationship. A possible source of systematic variance within groups was the speed of acquisition. A scatter plot of acquisition versus extinction slopes revealed a clear relationship: the steeper the acquisition slope, the steeper the extinction slope.

Therefore, an analysis of covariance was performed on the *slopes* of the extinction curves, correcting for the slopes of the acquisition curves. The acquisition slopes were estimated from the first three blocks of trials and the extinction slopes from the last five blocks of trials. The results of the analysis of covariance are presented in Table 3. Since slopes were used, there is no Blocks of Trials effect. Thus, the Blocks × Groups effect of Table 2 becomes Groups effect; Blocks × Populations becomes Populations effect; and Blocks × Groups × Populations becomes Groups × Populations.

The Groups effect is not significant indicating that with acquisition slope controlled the Cigarettes-Candy group does not extinguish faster than the "Good" group. The Groups × Populations interaction was also nonsignificant, which confirms what the analysis of variance suggested: the differences in the shape of the four extinction curves in Fig. 1 are random. Thus, with statistical control of acquisition slope, only one significant effect remains. The Populations variance is significant, indicating that college *S*s extinguish significantly faster than psychiatric patients.

Verbalization.—At the end of the experiment each *S* was queried about the task and the relationship between *S*'s responses and *E*'s behavior. Some *S*s were able to state that they were given chips (or *E* said "Good")

whenever they used first-person pronouns; these were the verbalizers. Other *Ss* offered either mistaken explanations or none at all (nonverbalizers). Again, only the data of Groups I and II will be analyzed.

The number of verbalizers in each group of 26 *Ss* was as follows: Group I, college—11; Group I, patient—6; Group II, college—22; and Group II, patient—10. The larger number of verbalizers in the Cigarette-Candy groups should be regarded as an artifact. The *Ss* in these groups were told at the beginning of acquisition that on some trials they would receive chips (to be turned in for cigarettes or candy). The *Ss* in the "Good" groups received no comparable instruction and therefore were less aware of the contingency between the reinforcement and their responses. The presence of more college verbalizers than patient verbalizers may be attributed to the greater alertness and verbal ability of the students. Interestingly, the presence of more college verbalizers did not lead to steeper acquisition curves by the college *Ss* (cf. Fig. 1). Similarly, the acquisition curves of Group II are not steeper than those of Group I, despite the larger number of verbalizers in Group II.

However, when each of the four groups is divided into verbalizers and nonverbalizers, the picture changes. Table 4 contains the acquisition data relevant to verbalization. With college *Ss* the trends in Group I and Group II are the same. The frequency of first-person responses increases over blocks of trials for the verbalizers, but the nonverbalizers manifest virtually no learning. The same trend is present in Group II patients: verbalizers learn, and nonverbalizers do not. However, the data of Group I patients are discrepant; here both verbalizers and nonverbalizers manifest learning. In fact the increment from the first block to the last block of trials is approximately the same for verbalizers and nonverbalizers.

TABLE 4

MEAN FREQUENCY OF FIRST-PERSON RESPONSES DURING ACQUISITION FOR VERBALIZERS AND NONVERBALIZERS

Blocks of 20 Trials	Group I		Group II	
	Verb	Nonverb	Verb	Nonverb
College				
1	9.4	7.5	7.7	9.0
2	12.3	8.1	12.5	7.0
3	15.0	7.9	15.0	5.5
4	16.5	8.9	15.5	7.8
Patients				
1	7.8	9.5	5.9	10.8
2	7.0	11.2	10.8	11.0
3	11.3	13.5	14.1	11.8
4	12.7	14.6	14.3	11.4

DISCUSSION

The extinction data yielded two major findings. First, with acquisition slope controlled Cigarettes-Candy did not yield faster extinction than "Good." Second, with acquisition slope controlled college students extinguished significantly faster than patients, regardless of the kind of reinforcement received during acquisition. This finding fits in with the generally accepted notion that psychiatric patients have more difficulty in unlearning and relearning.

Previous studies of verbal conditioning and extinction have not used cigarettes or candy as a reinforcer. Therefore in attempting to integrate the present results with those of other studies only the data of the "Good" groups will be used. In order to facilitate comparisons, extinction data from the studies of Cohen et al. (1), Grossberg (2), and Klein (4) have been plotted in Fig. 2,

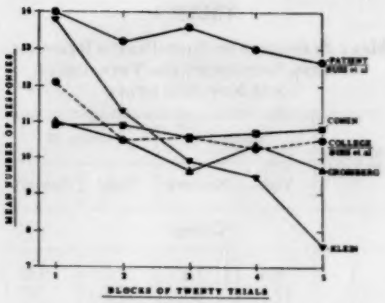

FIG. 2. First-person responses for last 20 acquisition trials and extinction trials by *Ss* in various studies.

together with the present extinction data. In this graph the initial point represents the last block of acquisition trials.

In Fig. 2 three curves appear to cluster together, those of Cohen et al., Grossberg, and the present college group. The *Ss* in these three groups were all normals, Cohen's being medical patients and Grossberg's being college students. None of these three curves shows appreciable extinction, suggesting that for normal *Ss* "Good" leads to extremely slow extinction.

The other two curves, Klein's and that of the patients of the present study, start the extinction series at roughly the same height. Both groups consisted of psychiatric patients, and in fact many of Klein's *Ss* were drawn from the same hospital as the present *Ss*. The similarity between the two curves ends at the initial point. Klein's curve drops sharply, while that of the present study remains relatively flat.

Klein's curve is the only one in Fig. 2 with a clearly negative slope; the others are relatively flat. A possible source of this discrepancy is speed of acquisition. In acquisition Klein's *Ss* initially made 6 first-person responses out of 20 and ended by making roughly 14 first-person responses out of 20. While they arrived at the same final point of acquisition as did the patient *Ss* of the present study, the patient *Ss* of the present study started by making roughly 9.5 responses out of 20. Thus, Klein's acquisition slope was somewhat steeper than that of the patients in the present study. In fact, Klein's acquisition curve is steeper than the acquisition curves of the other four comparable groups. Since acquisition slope was found to be important in the present experiment, it seems reasonable to attribute Klein's steeper extinction curve to faster conditioning.

Two explanations have been mentioned for the very slow extinction that usually follows "Good." First, secondary reinforcement is operating during extinction and serves to maintain the previously acquired response. Second, withholding a verbal reinforcer does not lead to a decrement in response strength such as found when a nonverbal reinforcer is withheld. The analysis outlined above suggests that a third explanation is more reasonable: the reinforcer "Good" usually leads to slow acquisition, and this slow acquisition in turn leads to slow extinction.

A unique aspect of the experimental situation employed here is the role of *E*. The *E* delivers the reinforcement "Good," and it seems likely that different *E*'s vary in their manner of saying "Good." The vocal inflections and nonverbal behavior of *E* in this rather close interaction with *S* probably exert considerable influence on the latter's learning. Such variations in *E*'s behavior may account for the differences between Klein's study and the others reviewed here. It would be interesting to determine the effect of removing *E* from the situation and having the reinforcement "Good" delivered by a phonograph recording.

The final issue raised by the present experiment concerns verbalization and learning. Again only the data of the "Good" groups will be discussed, since in the Cigarette-Candy groups verbalization was aided by the instructions. In this study 11 out of 26 college *Ss* verbalized, as did 6 out of 26 patient *Ss*. Grossberg (2) reported that 5 out of 15 college *Ss* verbalized, but Klein (4) and

Cohen et al. (1) both reported that virtually none of their *S*s were aware of the responses that were being reinforced. The discrepancy in number of verbalizers is probably due to different questions being asked. Grossberg's questions included "What was going on?", and in the present study each *S* was asked if he knew the connection between what he did and what *E* did. Such active interrogation undoubtedly uncovered more correct verbalizations than did the more circumspect questions employed by Klein and Cohen et al. Much has been made of verbal conditioning without awareness. Before an investigator concludes that he has demonstrated such conditioning, he should insure that his questioning is rigorous enough to exclude the possibility that *S* did not understand what was being asked.

Summary

This study investigated the effect of the reinforcers "Good," Cigarettes-Candy, and Poker Chips on operant verbal conditioning and extinction. The task was to select one of six pronouns plus the verb on a 3 × 5-in. card and make up a sentence. Each card constituted a trial, and there were 20 free responding trials, 60 acquisition trials, and 80 extinction trials. A total of 156 *S*s were used, including psychiatric patients and college students, men and women.

It was found that Poker Chips were ineffective as a reinforcer but that "Good" and Cigarettes-Candy were effective. Acquisition slope was found to be related to extinction slope. With acquisition slope controlled statistically, it was found that college students extinguished significantly faster than psychiatric patients.

The nonverbal reinforcer (Cigarettes-Candy) did not yield significantly faster extinction than the verbal reinforcer ("Good"). The flatness of extinction curves in this and previous studies was attributed to slow acquisition. Finally, the relationship between awareness and learning was found to be complex, depending in part on the population used and in part on the nature of the questions used to determine awareness.

References

1. Cohen, B. D., Kalish, H. I., Thurston, J. R., & Cohen, E. Experimental manipulation of verbal behavior. *J. exp. Psychol.*, 1954, **47**, 106–110.
2. Grossberg, J. M. The effect of reinforcement schedule and response class on verbal conditioning. Unpublished doctor's dissertation, Indiana Univer., 1955.
3. Hartman, C. L. Verbal behavior of schizophrenic and normal subjects as a function of types of social reinforcement. Unpublished doctor's dissertation, State Univer. of Iowa, 1955.
4. Klein, S. Conditioning and extinction of operant verbal behavior in neuropsychiatric hospital patients. Unpublished doctor's dissertation, Indiana Univer., 1954.
5. Lindquist, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1953.
6. Taffel, C. Anxiety and conditioning of verbal behavior. *J. abnorm. soc. Psychol.*, 1955, **51**, 496–501.

(Received September 20, 1957)

Journal of Experimental Psychology
Vol. 56, No. 2, 1958

MEANINGFULNESS OF MATERIAL, DISTRIBUTION OF PRACTICE, AND SERIAL-POSITION CURVES

HARRY W. BRAUN AND SYDNEY P. HEYMANN

University of Pittsburgh

Of the various theories that have been advanced to explain the bow-shaped serial-position curves which are typical in serial verbal learning, the most elaborate and perhaps best substantiated is that of Lepley (3) and Hull (2). According to this theory, remote forward associations (remote excitatory tendencies) are analogues of trace conditioned responses, and the remote forward associations between items in a list are inhibited by a process comparable to inhibition of delay. Furthermore, the amount of inhibition at each serial position is proportional to the number of remote excitatory tendencies spanning that position. Since these tendencies are most numerous in the middle of a series, inhibition is concentrated there and, consequently, the greatest number of errors in learning. Following from this theory is the prediction that distributed practice, produced by lengthening interitem or intertrial interval, will result in a relative loss of inhibition, as manifested by a decrease in errors at each position, with the greatest decrease in the central portion of the list.

This prediction was tested by Hovland (1) who had four groups of *Ss* learn the same list of 14 nonsense syllables under four practice conditions: (*a*) 2 sec. between items and 6 sec. between trials, (*b*) 2-sec. rate and 2 min. 6 sec. between trials, (*c*) 4-sec. rate and 6 sec. between trials, and (*d*) 4-sec. rate and 2 min. 6 sec. between trials. Distributed practice resulted in fewer errors than massed practice, but interitem distribution was more effective than intertrial distribution. The effect of distributed practice was greatest in the middle of the series as predicted by inhibition theory.

Recently, McCrary and Hunter (4) questioned the appropriateness of the inhibition theory to explain the serial-position effect. When they replotted Hovland's curves in terms of percentage of total errors at each position, the curves for the different conditions of learning were essentially similar. Thus, the reduction of errors by distributed practice occurs throughout the list in proportion to the total errors and in the same manner as in massed practice. McCrary and Hunter then had *Ss* learn a list of 14 nonsense syllables and a list of 14 names familiar to the *Ss*. Curves of mean absolute errors at each position for both classes of materials were bow-shaped, but the gradient of serial-position effect for the nonsense syllables was steeper than that for the names. The curves of percentage plots, however, were not different. Similar results were obtained for mean errors and percentage plots when the curves for the fast and slow learners in that study were constructed. Since these results were produced by the manipulation of meaningfulness of material and learning ability of *S* and not by variation of interitem and intertrail intervals, McCrary and Hunter concluded that a multiple- rather than a single-factor theory is required to handle the bow-shaped curve of the serial-position effect.

The purpose of the present experiment was to show how the systematic variation of meaningfulness of material and of interitem and intertrial

TABLE 1

MEANS AND *SD*'S FOR EACH GROUP OF TRIALS TO CRITERION AND TOTAL ERRORS IN LEARNING

Group	Meaningfulness	Interitem Interval	Intertrial Interval	Trials to Criterion		Total Errors	
				Mean	*SD*	Mean	*SD*
I	High	2″	6″	19.83	9.37	128.33	60.29
II	High	4″	6″	7.83	2.48	37.17	14.80
III	High	2″	2′6″	19.33	8.29	120.08	48.91
IV	High	4″	2′6″	8.25	2.89	41.67	14.86
V	Low	2″	6″	37.67	17.80	246.00	111.35
VI	Low	4″	6″	26.50	10.29	168.00	65.86
VII	Low	2″	2′6″	25.67	10.81	182.33	93.06
VIII	Low	4″	2′6″	20.50	7.16	129.92	42.01

intervals affects both the absolute and relative distribution of serial-position errors. Although they varied familiarity of material, McCrary and Hunter did not manipulate distribution of practice. Hovland varied distribution of practice but not the association value of his nonsense syllables. The present study duplicated Hovland's conditions of massing and distributing both items and trials with serial lists of high and low meaningfulness.

METHOD

This investigation utilized a 2 × 2 × 2 factorial design. The main experimental variables were: meaningfulness of material in two variations—high and low; interitem interval in two conditions—2 sec. and 4 sec.; and intertrial interval in two conditions—6 sec. and 2 min. 6 sec. Thus, there were 8 groups of 12 *Ss* each. All *Ss* were undergraduate men and women at the University of Pittsburgh and all were naive to the formal rote-learning situation. A random order of the eight experimental conditions (groups) was made up. The first 12 *Ss* were assigned to the first condition (group), the second 12 *Ss* to the second condition (group) and so on until all 96 *Ss* were run. It was necessary to impose a 2 hr. maximum on the learning period since it was impossible to detain unpaid volunteers longer. Two *Ss* in Group VII did not reach the criterion of learning in this time. Their trials-and-error scores were recorded as those which existed at the end of the 2-hr. period.

The high- and low-meaningful lists of 12 items each were selected from Noble's *m*-scale (**5**) and were constructed following his rule (**6**). The mean *m*-values of the lists were 7.85 (high) and 1.29 (low). The high-*m* words were: captain, zebra, garment, dinner, village, money, insect, jewel, army, heaven, office, wagon. The low-*m* terms were: volvap, neglan, latuk, tarop, quipson, gokem, kupod, polef, zumap, nostaw, sagrole, bodkin. Material was presented on a Hull-type memory drum, the anticipation-pronunciation method was used, and a criterion of one perfect repetition prevailed. Each item appeared once in every ordinal position for each condition (group), but sequence remained invariant.

Color naming was interpolated in the longer intertrial interval. A color chart which contained 30 small squares of five colors repeated in random order lay on the table in front of *S*. At a signal from *E*, *S* named these colors at a constant rate of about 1 sec. The *E* also signaled the end of the color naming.

RESULTS

Trials to criterion.—The mean trials to criterion and the *SD*'s of the eight distributions appear in Table 1. For both high- and low-meaningful lists, distribution of practice, effected by increasing either interitem or intertrial intervals, was associated with fewer trials to learn.

Absolute errors.—Figure 1 shows mean errors by serial-position curves made by the four groups who learned the high-meaningful list. It is obvious that interitem distribution is as-

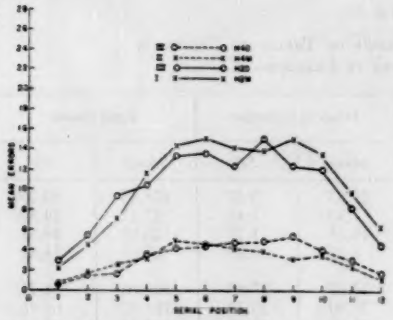

FIG. 1. Mean error serial-position curves for the high-meaningful list.

sociated with fewer errors. Furthermore, with each of the interitem intervals, variation of the intertrial interval does not affect over-all performance. In contrast with this neat state of affairs, the serial-position curves for the low-meaningful list show greater variability as demonstrated in Fig. 2. Again, however, it appears that interitem distribution results in more efficient learning. Especially noteworthy is the finding that with low-meaningful material the longer intertrial interval is related to superior performance over the shorter intertrial interval of 6 sec.

The means and *SD*'s of total errors for the eight groups appear in Table 1

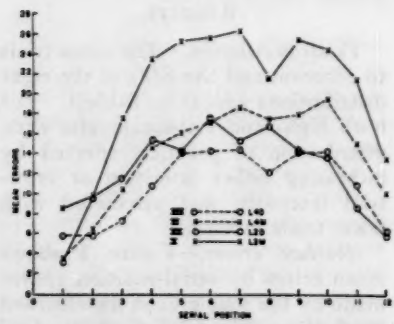

FIG. 2. Mean error serial-position curves for the low-meaningful list.

and the results of an analysis of variance of absolute error scores in Table 2. The main effects of meaningfulness and interitem interval are significant beyond the .01 level, but intertrial interval is significant at approximately only the .01 level. Again, Hovland's (1) findings are confirmed. Serial position is, of course, a significant source of variation as are the Serial Position and Meaningfulness interaction, the Serial Position and Interitem Interval interaction, and the Serial Position and Groups interaction. None of the other simple or higher order interactions is significant.

The use of *t* tests showed the same pattern of differences in absolute errors among the four high-*m* groups (I through IV) as in log trials to criterion. All comparisons of the four low-*m* groups (V through VIII) on absolute errors showed significant

TABLE 2

ANALYSIS OF VARIANCE OF RAW ERRORS IN LEARNING

Source	*df*	*MS*	*F*
Total	1151		
Between *Ss*	95		
Between groups	7	4876.85	12.762**
Meaningfulness (*m*)	1	19,900.12	52.077**
Interitem interval (I)	1	11,250.00	29.440**
Intertrial interval (T)	1	1391.28	3.641
m × I	1	191.75	—
m × T	1	1200.50	3.142
I × T	1	183.68	—
m × I × T	1	20.60	—
Ss within groups	88	382.13	
Within *Ss*	1057		
Serial position (SP)	11	1424.09	51.898**
SP × *m*	11	121.88	4.441**
SP × I	11	118.53	4.319**
SP × T	11	34.00	1.239
SP × Groups	77	51.70	1.884*
Pooled triple interaction	44	21.87	
Residual	892	27.44	

* $P < .05$.
** $P < .01$.

mean differences, except that of groups VI and VII. Thus, with both intertrial intervals, lengthening of the interitem interval was associated with superior performance, but the most efficient learning in these four groups occurred under conditions of spaced interitem and intertrial intervals (Group VIII).

Relative errors.—In plotting relative error curves, the procedure of McCrary and Hunter (**4**) was used. Errors at each position were expressed as a percentage of the total errors made under a given practice condition. The resulting curves for the four high-*m* groups appear in Fig. 3 and those for the four low-*m* groups in Fig. 4. When plotted in this manner, the eight error curves manifest a striking similarity. Although the high-meaningful family of curves appears to be displaced above the low-meaningful family, differences among conditions of practice seem to be minimal when relative errors at each position, rather than mean absolute errors, are plotted.

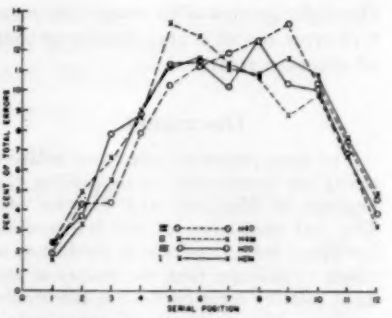

FIG. 3. Relative error serial-position curves for the high-meaningful list.

The results of an analysis of variance of percentage scores after an arc-sine transformation are shown in Table 3. The main effects of Meaningfulness and Interitem Interval are significant beyond the .01 level as is the Meaningfulness and Interitem Interval interaction. The Meaningfulness and Intertrial Interval interaction is significant at the .05 level. Serial Position is highly significant, but no other main effect or interaction is significant.

When *t* tests are applied, the pattern of significance of differences among

TABLE 3

ANALYSIS OF VARIANCE OF PERCENTAGES OF TOTAL ERRORS
(Arc Sine Transformation)

Source	*df*	*MS*	*F*
Total	1151		
Between *Ss*	95		
Between groups	7	32.76	9.867**
Meaningfulness (*m*)	1	105.85	31.882**
Interitem interval (I)	1	60.59	18.250**
Intertrial interval (T)	1	12.17	3.666
m × I	1	24.91	7.503**
m × T	1	14.62	4.404*
I × T	1	5.81	1.750
m × I × T	1	5.36	1.614
Ss within groups	88	3.32	
Within *Ss*	1057		
Serial position (SP)	11	1468.82	63.97**
SP × *m*	11	40.36	1.76
SP × I	11	20.66	—
SP × T	11	9.50	—
SP × Groups	77	20.51	—
Pooled triple interaction	44	18.26	
Residual	892	22.96	

* $P = .05$.
** $P = .01$.

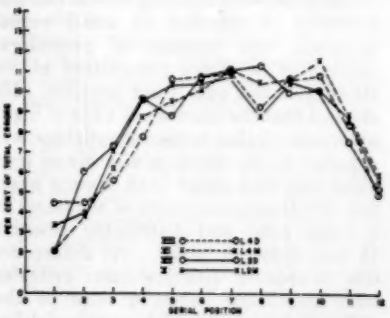

FIG. 4. Relative error serial-position curves for the low-meaningful list.

the eight groups with respect to relative error scores is very similar to that of absolute errors.

Discussion

The data presented above on relative errors are interpreted as supporting the findings of McCrary and Hunter (4). The fact that none of the interactions involving Serial Position is significant is taken to indicate that the shapes of the eight relative error curves are essentially similar. Manipulation of meaningfulness and of interitem interval may displace these curves upwards or downwards but it does not alter their form. It appears, then, that some factor (or factors) other than lapses in time is operating to decrease the number of inhibitory tendencies as predicted by the Lepley-Hull theory.

The finding that interitem distribution is relatively more effective than intertrial distribution in reducing errors in serial learning agrees with the results of Hovland (1) and Patten (7). Furthermore, the generality of this finding is enhanced in that it characterizes the learning not only of nonsense syllables but also of terms drawn from another and differently calibrated class of materials, i.e., Noble's *m*-scale (5). It will be recalled that this study employed the same interitem and intertrial intervals as did Hovland.

Tsao (8) has also investigated the relationship between meaningfulness and distribution of practice in serial verbal learning; but because of procedural differences, a direct comparison of his study and this one is not possible. He showed that the learning of a list of high-association-value nonsense syllables was superior to the learning of a list of low-value syllables under both massed practice (continuous exposure of the items at a 2-sec. rate) and distributed practice (1 min. between trials). All differences due to spacing with low-value syllables were significant, whereas none of the differences with the high-meaningful list was significant.

The finding that meaningfulness does not interact with either interitem or intertrial interval (in the analysis of raw error scores) is consistent with a similar result obtained by Underwood and Richardson (9) who used nonsense syllables which varied in association value.

Summary

This study was concerned with the relationship between meaningfulness of material, distribution of practice, and serial-position curves. High- and low-meaningful terms selected from Noble's *m*-scale were learned by the serial anticipation method with interitem intervals of 2 sec. and 4 sec. and intertrial intervals of 6 sec. and 2 min. 6 sec. Data were presented on trials to the criterion of learning, mean error serial-position curves, and relative (percentage) error serial-position curves. Meaningfulness and interitem interval were significant main effects in the three analyses. The data on relative errors were discussed in terms of their bearing on the Lepley-Hull theory of inhibition of delay.

References

1. Hovland, C. I. Experimental studies in rote-learning theory. III. Distribution of practice with varying speeds of syllable presentation. *J. exp. Psychol.*, 1938, **23**, 172–190.
2. Hull, C. L. The conflicting psychologies of learning—a way out. *Psychol. Rev.*, 1935, **42**, 491–516.
3. Lepley, W. M. Serial reactions considered as conditioned reactions. *Psychol. Monogr.*, 1934, **46**, No. 1 (Whole No. 205).
4. McCrary, J. W., & Hunter, W. S. Serial position curves in verbal learning. *Science*, 1953, **117**, 131–134.
5. Noble, C. E. An analysis of meaning. *Psychol. Rev.*, 1952, **59**, 421–430.
6. Noble, C. E. The role of stimulus meaning (*m*) in serial verbal learning. *J. exp. Psychol.*, 1952, **43**, 437–446.
7. Patten, E. F. The influence of distribution of repetitions on certain rote-learning phenomena. *J. Psychol.*, 1938, **5**, 359–374.
8. Tsao, J. C. Studies in spaced and massed learning. II. Meaningfulness of material and distribution of practice. *Quart. J. exp. Psychol.*, 1948, **1**, 79–84.
9. Underwood, B. J., & Richardson, J. Studies of distributed practice: XVIII. The influence of meaningfulness and intralist similarity of serial nonsense lists. *J. exp. Psychol.*, in press.

(Received September 23, 1957)

Journal of Experimental Psychology
Vol. 56, No. 2, 1958

ACQUISITION AND EXTINCTION OF A PARTIALLY REINFORCED RUNNING RESPONSE AT A 24-HOUR INTERTRIAL INTERVAL [1]

SOLOMON WEINSTOCK [2]

Indiana University

In order to account for the gross facts of acquisition and extinction, learning theorists have typically assumed that reinforced trials lead to increments and nonreinforced trials to decrements in habit "strength." One might intuitively expect that the larger the proportion of reinforced trials during acquisition the stronger the habit. This relationship, oddly enough, has been derived explicitly by only Bush and Mosteller (**3**, p. 89) for their linear operator learning model. Their model, as applied to the present situation, is mathematically identical with the models resulting from statistical learning theory (**5, 6**). Presumably a similar prediction may be obtained from Hullian theories. The present experiment is directed at testing this prediction.

In making a direct application of statistical learning theory to the runway situation, certain simplifying assumptions have to be introduced. This is equally true of Hullian theory and it would seem that Hull (**8**), at least implicitly, made these simplifications. For example, what would appear to be a chain of responses is treated as a single response. Further, the successive changes in stimulation are ignored and the stimulation is treated as a single stimulus event. The advantage in making these assumptions is that it is then possible to carry over to the runway a body of theory developed for classical conditioning and other similar situations. The price to be paid is the introduction of another source of possible theoretical error. Failure to verify theoretical predictions may now be due to the inadequacy of either the underlying theory or the simplifying assumptions.

A pilot study (**14**) for the present experiment, carried out in a runway and using a 24-hr. intertrial interval, yielded two major results: (*a*) an irregular ordering and no differences among terminal acquisition levels of groups receiving different proportions of reinforcement, and (*b*) slower extinction of the groups which received smaller proportions of reinforcement. As trials were very widely spaced, the extinction result was regarded as running counter to the expectations of the Hull-Sheffield hypothesis (**12**). The present study provides a further check on this result. The result of no differences in terminal acquisition levels is equivocal since the design of the pilot study did not allow the detection of small differences between groups. Use was made of the variance estimates from the pilot study to effect an increase in sensitivity through proper choice of the number of *S*s per group. In addition, further sensitivity was obtained through an extension of the range of the proportion of reinforcement variable.

The present experiment, then, was designed to explore parametrically the relationship between proportion of reinforcement and terminal acquisition level and to provide a statistically sensitive test of the presumed terminal differences between groups. It was also designed to differentiate between the discrimination model proposed by

[1] This paper is based on a Ph.D. thesis submitted to the Graduate School of Indiana University. The writer is very much indebted to W. K. Estes under whose direction the research was done. An abbreviated version of the present paper was given at the 1955 meeting of the Midwestern Psychological Association.

[2] Now at Lehigh University.

Bush and Mosteller (2) and one proposed by the writer (13), which differ mathematically only in Bush and Mosteller's use of a "discrimination operator."

METHOD

Subjects.—The *Ss* were 253 female albino rats, 90–120 days old at the start of the experiment. They were supplied by the Windsor Biological Gardens of Bloomington, Indiana, in one shipment. The *Ss* were assigned haphazardly to cages, five or six to a cage.

Apparatus.—An enclosed runway was formed by blocking off the left arm of a Y maze. The runway was 4 in. high and 4½ in. wide throughout. The starting box was 8 in. long. The running alley was 24 in. long and the arm of the Y which was used as a goal box was 18 in. long. The angle between the stem and arm of the Y was 30°. The runway was painted a flat gray and covered with hinge-mounted Plexiglas sections.

A 60-w. incandescent bulb was mounted in a diffusing globe 2½ ft. above the runway at a point directly above the juncture of the stem and the arm of the Y. Cheesecloth was draped 2 ft. above the runway in varying numbers of layers to provide an illumination level of approximately 2 ft.-candles throughout the runway.

A slot 2½ in. wide and 2 in. high was cut in the right side wall at the end of the arm of the Y and permitted access to a recessed drinking tube similar to those used in the regular watering periods. A flat gray slide could be inserted to prevent access to the drinking tube.

The starting box was separated from the running alley by a Plexiglas door which opened through a slot in the right wall. A sheet-metal door, painted flat gray, also opened through a slot in the right wall and separated the running alley from the arm of the Y.

Slots were cut in the side walls to permit the passage of light beams. The slots were placed 6, 18, and 30 in. from the starting box door. Three Standard Electric .01-sec. timers were activated when the starting box door was opened and deactivated by photocells when the appropriate light beam was interrupted.

The light source over the runway was activated when the lid of the starting box was closed and deactivated when the lid of the goal box was opened. When desired, a motor-driven commutator could be inserted into the light circuit to provide flicker at the rate of two on-off cycles per second.

The experiment was run in a lightproof room, separated from the animal colony by a curtain. A large exhaust fan in the animal colony ran continuously and provided a masking noise. Temperature in the colony room was maintained at 74° F. through the use of heating and cooling systems which were set to return the temperature to 74° when it deviated by a degree or more.

Procedure.—One continuously reinforced group, five partially reinforced groups, and five discrimination groups were run. The *Ss* were assigned randomly to groups. Reinforcements were assigned randomly by blocks of 12 trials so that the partially reinforced and discrimination groups received 10, 8, 6, 4, or 2 reinforcements per block of trials. The groups are referred to as Groups 12–0, 10–2, 8–4, 6–6, 4–8, or 2–10 according to the number of reinforced and nonreinforced trials in each block. Each of the continuously and partially reinforced groups received the same stimulation throughout the experiment, half receiving flicker and the other half receiving continuous illumination. Stimulation was correlated with reinforcement for the discrimination groups, half of each group receiving flicker and the other half receiving continuous illumination on reinforced trials.

The *Ss* were run under 22-hr. water deprivation at an intertrial interval of 24 hr. On reinforced trials, *Ss* had 30 sec. access to the drinking tube in the arm of the Y which served as a goal box. On nonreinforced trials, during both acquisition and extinction, *Ss* spent 30 sec. in the goal box prior to removal. The *Ss* were given unrestricted access to water for 1½ hr. beginning ½ hr. after the completion of their run.

The *Ss* were run in the same order each day. Just before its run, each *S* was removed from the home cage and brought into the darkened experimental room. She was then put in the starting box facing the starting box door. When the starting box lid was lowered, the stimulation *S* was to receive came on. Three seconds later *E* opened the starting box door, automatically starting the three timers. After interruption of the last photocell beam, a stop watch was started and the goal box door was closed. During the 30-sec. period *S* spent in the goal box, *E* recorded the readings on the three timers and reset the apparatus and clocks. When *S* was removed from the goal box, the illumination was automatically discontinued. The *S* was then returned to its home cage. As *E* returned *S* to its home cage, he set up the reinforcement and illumination conditions for the next *S* and reopened the door to the goal box.

The *Ss* were put on their watering schedules

one week prior to the beginning of the experiment and received access to water 1½ hr. daily. Each *S* was given 3–4 min. of gentling twice during the week preceding the experiment in addition to gentling during the assignment to cages and marking of the *S*s.

Continuous reinforcement was given to all groups for the first block of 12 trials. Intertrial intervals were greater than 24 hr. for Trials 2 through 5 because of slow running of *S*s. From Trial 5 to the end of the experiment it was possible to complete the daily session within 8 hr. On Trials 1 and 2 *S*s were given 20 min., and on Trials 3 through 12 they were given 5 min. to complete a run, and the same amount of time within which to begin drinking. Eight of the initial 253 *S*s were discarded for exceeding these limits more than once in the first 12 trials. On Trials 13 through 108, *S*s were reinforced according to the groups to which they had been assigned. A 2-min. period had been fixed for the completion of a run but no *S* had to be removed from the apparatus for exceeding this limit. On Trials 109 through 169 all *S*s received extinction, the discrimination *S*s receiving the same stimulation as they had received on reinforced trials during acquisition. A time limit of 5 min. was fixed for completion of a run during extinction but no *S* had to be removed from the apparatus for exceeding this limit. From Trials 13 through 169, *S*s were discarded only if they were completely unable to run owing to such factors as severe inner ear disease, paralysis, or an advanced stage of a neoplasm. Nine *S*s died or were discarded during acquisition and an additional 10 died during extinction.

It was desired to run all groups to asymptote during acquisition and extinction but, due to lack of advance information, no fixed number of trials could be built into the design of the experiment. A compromise solution was adopted. The decision to begin extinction was made on the basis of the first 72 trials at the end of which the groups appeared to have leveled off. Three additional blocks of 12 trials were to be run and statistical analyses for asymptotic behavior were to be made on the last 24 trials regardless of the nature of the data collected. A similar decision was made for extinction. On the basis of the first 30 trials it was decided to run another 30 trials and base asymptotic analyses on the last 20 trials.

Results

The latency and the two running-time measures yielded the same results and only those for the latency measure will be presented. To satisfy the requirement of homogeneity of variance in the analysis of variance, this measure was converted to reciprocal latency throughout.

Discrimination.—The mean reciprocal latency for each *S* in the discrimination groups on nonreinforced trials was subtracted from the mean reciprocal latency for reinforced trials for the blocks of trials from 85 to 108. If *S*s had formed a discrimination, this difference would have been positive. The mean of this measure for all five discrimination groups was equal to −0.01, indicating slightly more rapid running on nonreinforced than on reinforced trials. By analysis of variance the grand mean was found to be not significantly different from zero, while the differences among groups were also not significant at the 5% level.

Thus it is clear that the groups did not form a discrimination. They were not significantly different from the partially reinforced groups in all analyses of variance. Hence, in the later analyses each discrimination group was pooled with the partial reinforcement group receiving the same proportion of reinforcement. The pooling led to no changes in any of the statistical decisions made and was done exclusively to obtain more smoothness in the acquisition and extinction curves.

Acquisition.—In all analyses of variance for both acquisition and extinction the differences due to flicker and continuous illumination were not significant. Hence, the two halves of each group receiving different illumination were pooled for simple analyses of variance.

In Fig. 1 mean reciprocal latency is plotted by blocks of 6 trials from Trial 14 to the end of acquisition. The two groups receiving the smallest

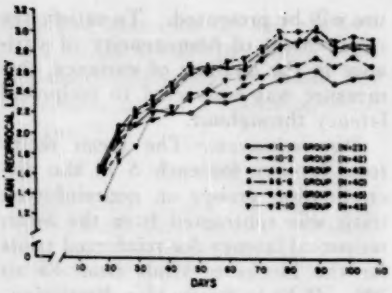

FIG. 1. Mean reciprocal latency as a function of trials by blocks of six for the acquisition phase of the experiment.

proportions of reinforcement ran most slowly for the first two blocks of 6 trials. An analysis of variance was made for the mean reciprocal latency for each *S* for Trials 14–25. The differences between groups are significant at the 5% level. Bartlett's test showed that the hypothesis of homogeneous variance could not be rejected. Hence, the differences are to be attributed to the differences between group means.

From Fig. 1 it may also be seen that Groups 12–0 and 10–2 ran most slowly for approximately the last 50 trials of acquisition. An analysis of variance for Trials 89–109 yielded differences between groups significant beyond the 1% level. The group means are presented in Table 1. Again Bartlett's test failed to reject the hypothesis of homogeneous variance.

A test for trend across the last two blocks of trials was made. The mean reciprocal latency for each *S* for the block of Trials 86–97 was subtracted from the corresponding mean for Trials 98–109. The mean difference here was −0.02, indicating slower running in the last block of trials. The *F* due to the mean difference was significant at the 5% level. The *t*'s for individual groups were not significant at the 5% level with the exception of Group 6–6 for which *t* was negative and significant at the 5% level. On the basis of this analysis and the forms of the curves in Fig. 1, it seems reasonable to assume that the groups were close to their asymptotic speeds.

Extinction.—Figure 2 shows that extinction proceeded in an orderly fashion with the smallest proportion of reinforcement groups extinguishing most slowly. Further, the groups appear to have leveled off during the last 20 extinction trials at different terminal levels. The 12–0 curve shows an unexpected upward trend after approximately the twentieth extinction trial. Apart from initial crossing-over of curves, attributable to different terminal acquisition levels, the only cross-overs that occur are between groups representing adjacent proportions of reinforcement.

The 60 extinction trials were divided into three blocks of 20 trials and mean reciprocal latency was computed for each *S* for each of the blocks. In addition, a mean reciprocal latency was taken for all 60 trials for each *S*. Four analyses of variance yielded *F*'s due to the proportion of reinforcement variable

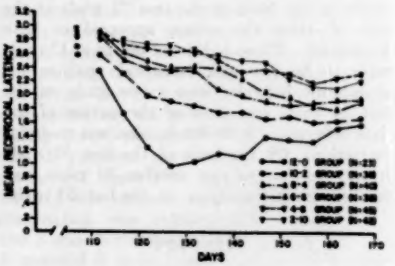

FIG 2. Mean reciprocal latency as a function of trials by blocks of five for the extinction phase of the experiment.

which were significant beyond the 1% level. However, Bartlett's test rejected the hypothesis of homogeneous variance except for the analysis for Trials 150–169. Hence, an unequivocal statement relating the significant F's to differences in group means rather than group variances cannot be made.

A modified t test proposed by Cochran and Cox and described by Edwards (**4**) was used to compare the two a priori extreme groups for each of the analyses and in all cases yielded a t significant beyond the 1% level.

Further, according to an empirical statistical study by Norton cited in Lindquist (**10**), the fact of nonhomogeneous variance should but slightly affect the conclusions to be drawn from an F test with respect to differences between group means under conditions which seem to be met by the present data. These considerations, coupled with the large mean differences observed, lead the writer to conclude that the significant F's may be attributed to differences both in group means and in group variances.

Since Groups 12–0 and 10–2 ran most slowly during acquisition and extinguished most rapidly, an analysis of covariance was run on the mean reciprocal latencies for each S for the 60 extinction trials with the mean reciprocal latency for Trials 86–109 as the control variable. The F's for the adjusted means are still significant beyond the 1% level.

A trend test was made for the last two blocks of 10 extinction trials. The statistical test was the same as that on the last two blocks of acquisition trials. The F for the grand mean of the differences between the two blocks was not significant at the 5% level. Since there is an independent interest in Group 12–0, a t test for that group indicated a significant trend ($P = .05$); the Ss ran more rapidly in the last 10 trials.

TABLE 1

INITIAL (L_0) AND ASYMPTOTIC (L_∞) LATENCIES AND SLOPE CONSTANTS (c) IN EXTINCTION AS A FUNCTION OF PROPORTION OF REINFORCEMENT

Group	$L_0(\bar{L}_{86-109})$	$L_\infty(\bar{L}_{150-169})$	c
12–0	2.59	1.57	.138 (.24[a])
10–2	2.73	1.64	.072
8–4	2.87	1.92	.056
6–6	2.91	1.84	.039
4–8	2.97	2.26	.050
2–10	2.86	2.16	.029

[a] The value .24 is obtained for c when the last 10 extinction trials are taken as asymptotic. This leads to a particularly poor fit for Group 12-0. A more sensible slope constant is obtained by neglecting the upward swing in the curve for Group 12-0 and by considering only the first 30 extinction trials. When this is done, under the assumption that Trials 21–30 are representative of asymptote, a value of .138 is obtained.

This series of statistical tests indicates that there are differences between groups during early, middle, and late extinction and that the groups level off at different extinction asymptotes as a function of proportion of reinforcement. The question arises whether the extinction differences are due not only to differing extinction asymptotes but also to differing rates, or slopes, of the extinction process. A crude answer to this question can be obtained by fitting a decay function

$$L_n = L_\infty - (L_\infty - L_0)e^{-cn} \quad [1]$$

to each group's extinction data by a least-squares criterion and getting an estimate for the slope constant c. In making the estimates, the value of the latency at the beginning of extinction, L_0, was taken as the mean for the last 24 acquisition trials and the asymptotic latency value, L_∞, was taken as the mean of the last 10 extinction trials. The values of c obtained are presented in Table 1,

from which it may be seen that, with one exception, the slope constants are related in an orderly way to proportion of reinforcement. Good fits to the data were obtained except in the case of Group 12–0 whose extinction curve appears to be **U** shaped. Estimation of a slope constant for each *S*, so that statistical tests may be used to determine the significance of the obtained differences, did not seem warranted by the present tentative state of extinction theory.

Discussion

The failure to obtain evidence of discrimination makes impossible any evaluation of the need for a theoretical "discrimination operator" under the present conditions. Animals run under conditions of 10 trials per day at approximately a 5-min. intertrial interval using the same discriminatory stimuli showed clear evidence of having formed the discrimination (**13**). This leads one to suspect that, at least for runway discrimination situations, trial spacing is a key variable.

The extinction data replicate those of the pilot study (**14**) to the extent of providing an orderly family of extinction curves with the smaller proportion of reinforcement groups extinguishing most slowly. Along with the apparent differences in the slopes of the extinction curves, there are two features of theoretical interest: the differences in terminal extinction levels and the upswing in the Group 12–0 extinction curve. Neither of these two results seems amenable to treatment by contemporary theories of extinction.

The acquisition data provide a confirmation of the relationship found by Notterman (**11**) to hold between proportion of reinforcement and terminal acquisition level. Notterman's study was primarily concerned with the problem of secondary reinforcement and left doubt about whether the continuous reinforcement group, which received fewer total trials than the partial groups, had reached asymptotic responding. In addition, there were a number of potentially confounding variables which rendered difficult clear attribution of the group differences to the proportion of reinforcement variable. The present data, while in a sense little more than a confirmation of Notterman's finding, do not appear to suffer from these difficulties.

While the differences between terminal acquisition levels are clear, it should be remembered that a statistically powerful design was required to detect them. The obtained differences are of the order of 10 to 15% and are not nearly as great as the extinction differences. A theory which predicts equal asymptotes for groups receiving different proportions of reinforcement would, therefore, provide a good first-order approximation to the data.

It has already been noted that in order to handle the runway situation theoretically, the simplification of regarding it as a single-stimulus, single-response situation has been adopted; i.e., the sequential and multiple nature of the stimuli and responses has been ignored. It is the view of the writer that the simplification is inadequate and that, therefore, judgment must be withheld on the adequacy of contemporary learning theories which predict the opposite of the present acquisition results.

Consideration of the runway as a sequence of responses where rigid constraint is imposed only on the last member of the sequence leads us to examine other experimental situations of the same type. When this is done it is recalled that many *S*s in the Guthrie and Horton (**7**) study showed response chain components which were not functional in obtaining release from the apparatus. For example, they report that one cat consistently touched the wall before moving the pole to obtain release. According to them, many of the responses that occurred in the response sequence on early trials, whether functional or not, remained in the response chain that became stable.

This raises the general problem of what happens to what may be called "accessory, nonfunctional movements" in a response sequence which leads to terminal reinforcement. The Guthrie and Horton study would lead one to expect that a reasonable number of the accessory, nonfunctional movements remain in the chain if reinforcement is given on every trial. The same conclusion follows from various reports of "superstitious" behavior.

The implication of the above analysis for the interpretation of the present acquisition data is that nonreinforced trials provide an opportunity for accessory, nonfunctional movements to drop out and thus enable the partial reinforcement groups to achieve faster running levels. That interspersed nonreinforced trials may serve to filter out accessory, nonfunctional movements is implied by a study by Antonitis (1) in which rats ran across an open area and poked their noses into a slot which ran the length of one wall to get reinforcement at the opposite end of the apparatus. Antonitis measured the variability in the terminal response during acquisition, extinction, and reacquisition and found a considerable decrease in reacquisition as compared with acquisition. The implication here is that a decrease in variability in terminal position reflects a decrease in the over-all path taken and may be attributed to the elimination of accessory, nonfunctional movements from the response sequence. A number of "mechanisms" are available to account for the "purification" of the response sequence; however, at this time it is unclear which of them are operating in the runway situation. A similar explanation is presented by Lauer and Estes (9) to account for the smaller asymptotic latencies they obtained with successive acquisitions and extinctions in a jumping situation.

One further point concerning the acquisition data is worth noting: namely, that in the early trials the speed of running is directly related to proportion of reinforcement. The change in the nature of the relationship from early to late acquisition implies that variable parameters will be required in the theoretical learning functions used to describe the partial reinforcement learning curves.

Summary

The present study was designed to detect anticipated small differences in acquisition asymptotes of groups receiving varying proportions of reinforcement.

Rats ($N = 253$) were run in an enclosed runway at a trial spacing of 24 hr. They were run under 22-hr. water deprivation and received 30 sec. access to a water tube on reinforced trials. They remained in the goal box for 30 sec. on nonreinforced trials. Six proportions of reinforcement were randomly assigned by blocks of 12 trials: 12–0, 10–2, 8–4, 6–6, 4–8, and 2–10. There were 109 acquisition trials and 60 extinction trials.

The major experimental finding is that Groups 12–0 and 10–2 ran most slowly in the last two blocks of 12 acquisition trials. In extinction, an orderly family of curves was obtained with the groups with smaller proportions of reinforcement showing both smaller slopes and higher terminal extinction asymptotes. The theoretical significance of the results was briefly discussed.

References

1. Antonitis, J. J. Response variability in the white rat during conditioning, extinction, and reconditioning. *J. exp. Psychol.*, 1951, **42**, 273–291.
2. Bush, R. R., & Mosteller, F. A model for stimulus generalization and discrimination. *Psychol. Rev.*, 1951, **58**, 413–423.
3. Bush, R. R., & Mosteller, F. *Stochastic models for learning.* New York: Wiley, 1955.
4. Edwards, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.
5. Estes, W. K. Toward a statistical theory of learning. *Psychol. Rev.*, 1950, **57**, 94–107.
6. Estes, W. K., & Burke, C. J. A theory of stimulus variability in learning. *Psychol. Rev.*, 1953, **60**, 276–286.
7. Guthrie, E. R., & Horton, G. P. *Cats in a puzzle box.* New York: Rinehart, 1946.
8. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.

9. Lauer, D. W., & Estes, W. K. Successive acquisitions and extinctions of a jumping habit in relation to schedule of reinforcement. *J. comp. physiol. Psychol.*, 1955, **48**, 8–13.
10. Lindquist, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1953.
11. Notterman, J. M. A study of some relations among aperiodic reinforcement, discrimination training, and secondary reinforcement. *J. exp. Psychol.*, 1951, **42**, 273–291.
12. Sheffield, V. F. Extinction as a function of partial reinforcement and distribution of practice. *J. exp. Psychol.*, 1949, **39**, 511–526.
13. Weinstock, S. Acquisition and extinction of a partially reinforced running response at a 24 hour intertrial interval. Unpublished doctor's dissertation, Indiana Univer., 1953.
14. Weinstock, S. Resistance to extinction of a running response following partial reinforcement under widely spaced trials. *J. comp. physiol. Psychol.*, 1954, **47**, 318–322.

(Received September 30, 1957)

Journal of Experimental Psychology
Vol. 56, No. 2, 1958

EFFECTS OF SHOCK–INDUCED STRESS ON VERBAL PERFORMANCE [1]

W. DEAN CHILES

Aero Medical Laboratory, Wright-Patterson Air Force Base

Many problems are encountered in trying to interpret the results of experiments concerned with measuring the effects of stressful environments on performance. Of these problems, the one most difficult to resolve is that which deals with the role played by motivation in such experiments. This same obstacle is encountered when one tries to extend basic psychological theory to account for complex human behavior. The close similarity of these two problem areas indicates that an appreciable theoretical as well as methodological gain would accrue from a successful attack on the more basic of the two questions, i.e., the role of motivation in complex human behavior.

A number of recent psychological investigations have been concerned with the use of a psychological test, the Manifest Anxiety Scale (**6, 7**), as a means of assessing the general motivational level of the individual. The Manifest Anxiety Scale (MAS) was devised originally, and is now being used, because it was felt, first of all, that the test provided a measure of the individual's "built-in level of anxiety," and second, that this generalized level of anxiety should be related to the individual's drive level.

In the typical study involving the use of the MAS, *E* selects two groups of *S*s—one group scoring high and one scoring low on the test. Having thus constituted his groups, *E* requires *S*s to perform, for example, a verbal learning task. On the basis of his particular theoretical orientation, he then predicts the relative performance levels which the two groups would exhibit had they differed with respect to drive level. If the high-anxious group behaves in the same manner as the hypothetical high-drive group would behave, and the low-anxious group behaves as would the low-drive group, *E* assumes that he has confirmed the hypothesis that drive level has some specified effect on performance.

Within the framework of the Hullian behavior theory, anxiety is assigned the role of an irrelevant drive. Since in the Hullian theory effective drive is a generalized concept, if all other factors are held constant, the level of drive should determine the level of performance in a given situation, and this should be true without regard to the source of any irrelevant drives present.

Thus, if one were to duplicate an experiment which involved the use of the Manifest Anxiety Scale, but were to induce the differences in drive level through the introduction of a stressful stimulus, one should be able to repeat the essential aspects of the results of that MAS study.

Ramond (**1**) carried out a study in which two groups of *S*s, which differed with respect to anxiety level, were required to learn a list of verbal, stimulus-response relationships. These relationships may be characterized in terms of three factors: (*a*) Some of the relationships involved initially strong associations, and some involved initially weak associations; (*b*) the strong- and weak-association responses were in competition with each other; and (*c*) the weak-association

[1] The main findings of this experiment were presented at the 1957 meeting of the American Psychological Association.

relationships involved learning to respond with a word which was unrelated to the stimulus in preference to a second possible response which was a synonym of that stimulus.

On the assumption that the MAS score was indicative of the *S*'s drive level, Ramond predicted that the anxious *S*s would be superior on the strong-association items, but would be inferior on the weak-association items. He based his prediction on the fact that the higher drive level of the anxious group would give that group an advantage in response strength on the synonym-correct items; but on the weak-association items, the anxious group would have a relatively stronger response to extinguish before the initially weak response could become dominant.

The second part of Ramond's prediction was realized—that is, the nonanxious *S*s were superior on the weak-association items. But the first part of the prediction went awry; the anxious group was actually slightly inferior to the nonanxious group on the strong-association items. This unpredicted result Ramond explains in terms of generalization, i.e., the anxious *S*s, having a higher drive level, would experience more interference between the learning of the two lists than would the nonanxious group.

The purpose of the present experiment was to duplicate the study carried out by Ramond, except that the difference in drive levels would be induced experimentally rather than through the selection of groups on the basis of the MAS. Two ends were thus being sought. One had to do with the use of an electric shock as a means of inducing psychological stress and, hence, a higher drive level in a group of *S*s. The other had to do with the use of "experimental anxiety" as compared to Ramond's use of "test anxiety."

Method

The basic experimental procedure duplicated that used by Ramond with the following exceptions: (*a*) The *S*s used in the present study were not selected on the basis of their scores on the MAS. (*b*) The difference in the treatments of the two groups of *S*s was that one group responded to shock and the other to a buzzer in connection with an auxiliary reaction-time task; however, it should be noted that, in terms of the instructions, *S*s were led to believe that the reaction-time test was the primary task. (*c*) The actual word lists used, though prepared according to the scheme which Ramond followed, were taken from a standard copy of Roget's *Thesaurus*, since Ramond's word lists were not available. (*d*) The *S*s in the present experiment were given 48 rather than 32 trials on the test list, although the statistical analysis was applied in the same manner as was Ramond's.

Subjects.—The *S*s used were paid volunteer students from the University of Dayton. They were selected from the male subject pool on the basis of their willingness to be subjected to a harmless electric shock. The *S*s were placed unsystematically in one of the two groups. The *S*s in the buzzer group were told at the beginning of the experimental session that they would not be shocked at all during the experiment.

Apparatus.—The verbal task was presented using a memory drum with the following modifications: (*a*) The shutter mechanism was removed since choice rather than anticipation of the correct response was to be employed, and (*b*) the 8-sec. revolution period was divided into three equal parts to yield an exposure interval of 2.4 sec. and an interexposure interval of .25 sec.; there was a 5.3-sec. interval between successive lists.

The memory drum was used to present simultaneously three two-syllable adjectives, one on the left and two on the right. Four different orders of a list of 16 such presentations were typed in lower-case letters on an endless cloth tape. Parallel to these lists was typed a practice list of two-syllable nouns. The test materials, as were Ramond's, were selected so that one member of a pair of response words was a synonym of the corresponding stimulus word, whereas the other member of the pair was unrelated to the attached stimulus word. The following example will illustrate how the stimulus-response combinations were interrelated:

A	tranquil	serene	s	r_1
		rugged		r_2
B	peaceful	serene	s^1	r_1
		rugged		r_2

The presentation of A and of B would be separated by the presentation of one or more other stimulus-response combinations. Thus, *S* saw the same pair of response words twice in any given order of the list. The particular response word which was correct for a given one of the two stimulus words involved was predetermined on an arbitrary basis. For example, on Presentation A, *rugged* was the correct response to *tranquil*; therefore, on Presentation B, *serene*

was the correct response to *peaceful*. The combination, *tranquil-rugged*, was assumed to have a lower associative strength than the combination *peaceful-serene*. This method of presentation avoided serial learning of responses and required association of the response words with their particular stimuli.

The apparatus for the auxiliary task consisted of a "start" switch which closed a locking circuit, a buzzer or a constant current stimulator (depending upon the group in which the *S* was placed), and a switch controlled by *S* which broke the locking circuit. The *S*s in each group were required to keep a palm switch depressed at all times. The buzzer group operated the switch when the buzzer sounded during the 5.3-sec. interlist interval and the shock group in response to an electric shock applied through silver electrodes attached to the ankle. Since the presentation of a list took about 50 sec., *S*s in the shock group were shocked approximately once per minute throughout the course of the experiment. The buzzer was set at an intermediate loudness and was mounted on sponge rubber. The shock apparatus was set to deliver a constant current of .5 milliamp. through electrodes spaced approximately 1 in. apart.

Procedure.—The toss of a coin determined the group into which each *S* would be placed. This resulted in 30 *S*s in the shock group and 24 *S*s in the buzzer group. At the beginning, *S* was told that the primary purpose of the experiment was to determine the relative effectiveness of an electric shock in eliciting a response from an individual who was engaged in a rather absorbing mental task. The electrodes were then attached to the ankle of *S* if he were in the shock group but not if he were in the buzzer group.

The *S* was told that he would see a succession of presentations of "one stimulus word on the left and two response words on the right" and that his task was to "learn in each case which of the two words on the right is the correct response to make." The *S* stated aloud his choice at each presentation, learning from *E*'s reply—"Right" if correct and no answer if incorrect—whether a given answer was right or wrong.

The session began with 16 trials on the practice list. These practice trials were followed by 2 min. of rest. The *S* then went through three 16-trial sessions which were separated by 2-min. rest intervals during which *E* engaged *S* in conversation in order to prevent review of the materials. In each case a trial is defined as one presentation of the list of 16 S-R combinations. The shocks and buzzer soundings occurred throughout the practice and the training trials. The total time required was about 1.5 hr. per *S*.

Results

Figure 1 presents the data on the performance of the buzzer group on the strong as compared to the weak items. This comparison indicates that the initial strengths of the associated responses were in fact stronger than those of the nonassociated responses; the difference on which this inference is based is also apparent throughout the first 32 trials with a suggestion of that difference during the final 16 trials. A Wilcoxon *T* test, which compared the total number of correct responses under the two conditions, showed the difference over the first 32 trials to be significant at the .02 level. The difference over the final 16 trials was not significant.

Figure 2, which compares the performance of the shock group on the associated items with its performance on the nonassociated items, reveals the same relationship that is displayed by the buzzer group; i.e., the shock group performed better on the strong than on the weak items throughout the course of the experiment. The Wilcoxon *T* for the difference was significant ($P < .01$) both for the first 32 trials and for the last 16.

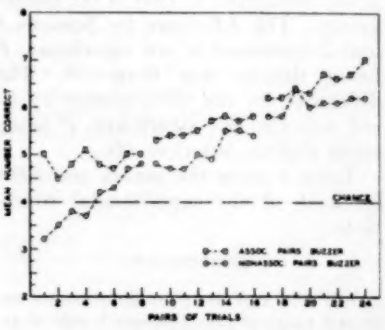

FIG. 1. A comparison of the performance on the associated vs. the nonassociated pairs by the buzzer group.

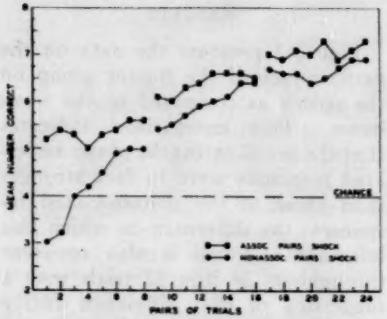

FIG. 2. A comparison of the performance on the associated vs. the nonassociated pairs by the shock group.

Figure 3 presents a comparison of the shock and the buzzer groups on the strong-association items. With the exception of one aberrant point (Trials 7 and 8) the shock group is superior to the buzzer group. The differences were significant, according to Mann-Whitney two-tailed U tests, as follows: for Sessions 1 and 2 combined, $P < .02$; and for Session 3, $P < .001$.

Figure 4 presents a comparison between the shock and buzzer groups on the weak-association items. With the exception of the first two pairs of trials, the performance of the shock group is superior to that of the buzzer group. The difference for Sessions 1 and 2 combined is not significant, P being slightly less than .09. The difference for the third session by itself was also not significant, P being again slightly less than .09.

Table 1 gives the means and SD's for each of the experimental conditions.

DISCUSSION

The attempt to provide stimulus-response pairs of two different levels of associative strength was successful. Therefore, for theoretical purposes, we are justified in speaking of differences in initial habit strength within the context described by Ramond.

It was hypothesized that the administration of an electric shock would produce an increase in the effective drive level of *S*s in the shock group. This being the case, the theory leads to certain predictions concerning the performance of the two groups.

First of all, in the case of the strong-associated items, the theory predicts that the group with the higher drive level will perform better on such items than will the low-drive group. The significant difference in favor of the shock group on these items bears out that prediction.

In the case of the nonassociated items, the theory specifies that the high-drive group will exhibit inferior performance until the point is reached at which the relative strengths of the two responses reverse, i.e., the correct response, through reinforcement, becomes the stronger of the two; when the correct response does become the stronger, the high-drive group will be superior. Again, as is seen in Fig. 4, the theoretical prediction is borne out. An analysis of variance showed that the salient statistic, the simple interaction variance between Trials and Groups for the nonassociated items, was significant at the .05 level; however, since the satisfaction of the assumptions necessary to this analysis could not be rigorously established, this test cannot be treated as being conclusive.

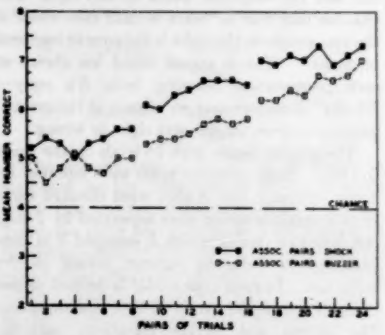

FIG. 3. A comparison of the shock and the buzzer groups on the associated pairs.

The original intention in adding the third session was to provide opportunity for the high-drive group to "catch up" to the low-drive group if the data were to have reflected the same type of result found by Ramond. Since the predicted relationships established themselves earlier in the training, there is no particular theoretical importance to the data for the additional 16 trials.

It will be recalled that Ramond's high-drive (anxious) group was inferior on both the associated and the nonassociated items—significantly so in the latter case. Thus, in two very similar situations, increased motivation was found to have essentially diametrically opposed effects. Inasmuch as Ramond attempted to explain his results in terms of the Hullian behavior theory, let us examine his findings with respect to the appropriateness of that interpretation. As was pointed out above, the theory yields a clear-cut prediction that the high-drive group will be superior on the associated items and, after an initial period of inferiority, on the nonassociated items as well. Therefore, in the case of the nonassociated items, we must conclude that Ramond was in error in interpreting his results as being in support of the theory. Ramond refers to a slightly different aspect of the theory in accounting for the results on the associated items. The slight inferiority of the high-drive group on these items is said to be a result of the fact that, as the nonassociated S-R connections are learned, generalization to the associated items takes place; therefore, because of their higher drive level, the anxious *Ss* experience more interference with the learning of the associated items

TABLE 1

TOTAL NUMBER OF CORRECT RESPONSES

Type of Response	Buzzer		Shock	
	Mean	*SD*	Mean	*SD*
Associated	167	28.7	188	35.0
Nonassociated	145	34.6	163	26.5

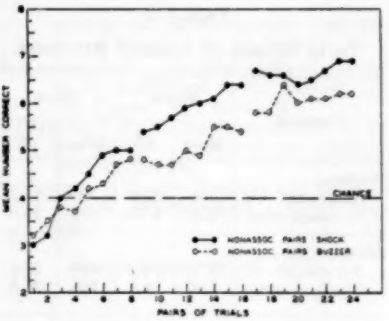

FIG. 4. A comparison of the shock and the buzzer groups on the nonassociated pairs.

than do the nonanxious *Ss*. The present experiment can be said to have shown this interpretation to be in error (though perhaps not conclusively), since there is no apparent theoretical reason that such generalization should not have occurred here as well as in Ramond's experiment.

Thus far, it has been tacitly assumed that a high MAS score can be treated as being related to a high level of drive. However, this assumption may be questioned.

In order to get further information pertinent to this question, subsequent to the completion of the present experiment, manifest anxiety scores were obtained on the 36 *Ss* who were available for testing. Fortunately, 18 were shock *Ss* and 18 were buzzer *Ss*. Each of these groups was subdivided at its median (a score of 14 in the case of the buzzer group and a score of 11 for the shock group) into high- and low-anxious subgroups. The performance of these subgroups in terms of the total number of correct responses is shown in Table 2, and an analysis of variance of these data is presented in Table 3. This analysis indicates that the effect of Anxiety is not significant ($F < 1$); the effect of Shock is significant ($.01 < P < .05$); and the effect of degree of Association is significant ($P = .005$). The interaction of Anxiety and Shock is not significant although P is less than .10. The other interaction F ratios are less than 1. The tests of primary interest are for the effect of shock, of anxiety, and of the interaction between these two. Considering only the buzzer group, we see that anxious *Ss* are slightly inferior on both types of items, which, it will be recalled, is the same result obtained by Ramond; in the

TABLE 2

TOTAL NUMBER OF CORRECT RESPONSES

Condition	Buzzer		Shock	
	Mean	*SD*	Mean	*SD*
Anxious				
Associated	166	26.6	200	26.7
Nonassociated	138	47.0	178	38.0
Nonanxious				
Associated	171	26.2	183	27.6
Nonassociated	159	24.5	156	26.8

present instance, however, the differences are undoubtedly not significant. The performance of the anxious *Ss* in the shock group presents a somewhat different picture. In this case, the anxious *Ss* are superior to the nonanxious *Ss* on both types of items.

The fact that the over-all effect of anxiety was not significant may have been a reflection of the relative magnitude of the Anxiety by Shock-buzzer interaction, i.e., when the effects of anxiety are averaged over the shock and buzzer conditions, any differences present tend to disappear.

With respect to the effects of shock, the results of the present experiment agree with those reported by Spence, Farber, and Taylor (3). In their experiment, which involved the effects of electric shocks on eyelid conditioning, they concluded that their data tend to support the position that so-called anxious *Ss* are best described as exhibiting a greater response to aversive stimuli than do nonanxious *Ss*. On the other hand, anxiety has been shown to increase the level of performance in simple, noncompetitional verbal learning situations which do not involve noxious stimuli (2, 4).

Thus in some situations, manifest anxiety produces results which show it to be related to drive (5); in others, the results show that increased drive results from the fact that anxious *Ss* react more to stressful stimuli than do nonanxious *Ss* (3); and, in still other situations, anxious *Ss* behave as though they actually had lower drive levels than nonanxious *Ss* (1, 2, 4).

TABLE 3

ANALYSIS OF VARIANCE OF ANXIETY DATA

Source	*df*	*MS*	*F*	*P*
Between *Ss*	35	1755		
Anxiety (An)	1	256	<1.	
Shock (S)	1	7980	5.28	<.05
An × S	1	4802	3.18	<.10
Error (*b*)	32	1512		
Within *Ss*	36	882		
Association (A)	1	8844	12.74	<.005
A × An	1	108	<1.	
A × S	1	85	<1.	
A × An × S	1	512	<1.	
Error (*w*)	32	694		
Total	71			

SUMMARY

In order to examine the effects of an experimentally manipulable irrelevant drive on verbal performance, 30 *Ss* were given 48 trials under a shock (high drive) condition, and 24 *Ss* were given the same number of trials under a buzzer (low drive) condition. The stimulus-response materials were of two types; one involved the learning of a synonym response which was in competition with an unrelated word and the other involved the learning of an unrelated word which was in competition with a synonym. The shock and the buzzer were presented in connection with a reaction-time task performed between trials (1 per minute) on the primary task. After the completion of the basic experiment, MAS scores were obtained on 18 shock *Ss* and 18 buzzer *Ss*.

The results of the basic experiment showed the shock group to be significantly superior on the associated items and, after being initially inferior, consistently superior on the nonassociated items. These results are in agreement with the theory which relates drive level to performance but they do not agree with those of Ramond. The breakdown of the data according to A scores showed the shock to have a considerably greater incremental effect on the high-anxious *Ss* than on the low-anxious *Ss*. Also, the anxious *Ss* under the buzzer condition were slightly inferior on both types of items, whereas under the shock condition, the anxious *Ss* were superior on both types of items.

REFERENCES

1. RAMOND, C. K. Anxiety and task as determiners of verbal performance. *J. exp. Psychol.*, 1953, **46**, 120–124.

2. SPENCE, K. W., FARBER, I. E., & MCFANN, H. H. The relation of anxiety (drive) level to performance in competitional and noncompetitional paired-associates learning. *J. exp. Psychol.*, 1956, **52**, 296–305.
3. SPENCE, K. W., FARBER, I. E., & TAYLOR, E. The relation of electric shock and anxiety to level of performance in eyelid conditioning. *J. exp. Psychol.*, 1954, **48**, 404–408.
4. SPENCE, K. W., TAYLOR, J., & KETCHEL, R. Anxiety (drive) level and degree of competition in paired-associates learning. *J. exp. Psychol.*, 1956, **52**, 306–310.
5. TAYLOR, J. A. The relationship of anxiety to the conditioned eyelid response. *J. exp. Psychol.*, 1951, **41**, 81–92.
6. TAYLOR, J. A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, **48**, 285–290.
7. TAYLOR, J. A. Drive theory and manifest anxiety. *Psychol. Bull.*, 1956, **53**, 303–320.

(Received September 26, 1957)

Journal of Experimental Psychology
Vol. 56, No. 2, 1958

ANXIETY LEVEL AND EGO INVOLVEMENT AS FACTORS IN CONCEPT FORMATION [1]

CONCETTA V. ROMANOW

Chicago Teachers College

This study deals with the acquisition of concepts associated with interfering responses of varying strength, as a function of two processes held to be indicators of drive level: manifest anxiety and ego involvement.

Research findings indicate that high-anxious *Ss* are superior to low-anxious *Ss* in acquiring simple CR's such as the eyeblink (**8, 9, 10, 12, 13**) and the GSR (**1, 5**). Other studies have shown that high-anxious *Ss* are superior when association value of words in a serial list is high and intralist similarity is low (**6**); when competition among responses in a paired-associates list is minimal (**11**); and, in problem solving, when the dominant set is congruent with the correct solution (**4**). On the other hand, high-anxious *Ss* are inferior to low-anxious *Ss* under the following conditions: in a serial list, when association value is low and intralist similarity is high (**6**); in learning paired associates, when the associative strength between stimulus and response is weak (**7**) and when competition among responses is strong (**11**); in problem solving, when the dominant set is not congruent with the correct solution (**4**).

These and similar findings indicate that high-anxious *Ss* are superior when tendencies competing with the correct response are weak, but they are inferior when there are strong tendencies interfering with the required response.

Spence and Taylor (**12**), assuming that manifest anxiety is an indicator of drive level, have proposed a theory based on Hull's concept of drive to account for such results. According to the Hullian formula $E = D \times H$, the strength of an excitatory tendency is a multiplicative function of drive and habit strength (**3**). According to this, increased drive would strengthen the excitatory tendencies of all habits in *S*'s hierarchy. However, when drive is increased, the excitatory potentials of strong habits in the hierarchy would be increased relatively more, as compared with the increase in excitatory potentials of weak habits. Consequently, when the correct response is of high habit strength, the high-drive *Ss* should perform better because weak competing responses are relatively weaker for them than for a low-drive group. Conversely, if the correct response is of weak habit strength and interfering tendencies are strong, then the high-drive *Ss* should be inferior to a low-drive group, since the strong competing responses are relatively stronger for them than for a low-drive group.

The present study involves a concept formation task in which the strength of the correct response in relation to competing responses is controlled and varied. Two separate experiments were conducted; in one, *Ss* differed in degree of manifest anxiety, while in the other, *Ss* differed in level of ego involvement. These parallel experiments were conducted under the assumption that both manifest anxiety and ego involvement are indicators of drive level. Thus, the same predictions based on Hullian theory were made for both studies: namely, that high-drive (high-anxious or highly involved) *Ss* would be superior when the correct response is strong and interfering tendencies are

[1] This study is based on a dissertation submitted to the Psychology Department of Northwestern University in partial fulfillment of requirements for a Ph.D. in psychology. The writer wishes to thank B. J. Underwood, director of the dissertation, for his encouragement and guidance.

weak, but that low-drive *S*s would be superior when the correct response is weak and interfering tendencies are strong.

The studies previously cited show that predictions from Hullian drive theory have been confirmed with respect to manifest anxiety; the inference is that differences in manifest anxiety reflect drive differences. This study also explores the effect of ego involvement when "correct" and interfering tendencies vary in strength in an attempt to apply Hullian theory to ego involvement. Since two separate experiments were performed, the present research does not provide information on the interaction of the two variables, anxiety and ego involvement.

Experiment I

Method

Design.—There were three groups of *S*s who differed in degree of manifest anxiety. The *S*s were required to respond with a descriptive adjective to each of 24 different nouns which were presented repeatedly in a serial list. These stimulus nouns formed six concepts of four nouns each, since one particular descriptive response was correct for four different nouns in the list. For two of these concepts (i.e., two "sets" of four nouns), the correct response was attached to the stimulus with high associative strength; for two other concepts, the correct response was one of moderate associative strength; and for the other two concepts, the correct response was one of low associative strength. Strength of tendencies competing with the correct response varied

TABLE 1

Noun Stimuli in the Experimental List and All Responses Given by Standardization *S*s With Frequency of 5% or Greater

(Correct responses are in italics)

Concept Dominance	Stimulus	Responses
High	Globe	*round* (95%)
	Wheel	*round* (94%)
	Spool	*round* (74%); small (9%); woody (7%)
	Doughnut	*round* (71%); sweet (7%)
	Dandelion	*yellow* (85%)
	Canary	*yellow* (82%); small (5%)
	Corn	*yellow* (81%)
	Mustard	*yellow* (75%); strong, sharp (14%)
Moderate	Gym	*big* (54%); smelly (21%)
	Walrus	*big* (46%); slimy (13%); black (13%)
	Gorilla	*big* (42%); hairy, furry (41%)
	Ocean	*big* (33%); blue (24%); wet (20%); deep (12%)
	Flannel	*soft* (54%); red (12%); itchy (8%); fuzzy (5%)
	Kitten	*soft* (41%); small (25%); hairy, furry (13%)
	Silk	smooth (41%); *soft* (39%); shiny (6%); slippery (5%)
	Bread	white (35%); *soft* (31%)
Low	Alley	dark (49%); *long* (16%); narrow (14%); dirty (10%)
	Eel	slimy (68%); *long* (15%)
	Asparagus	green (78%); *long* (9%)
	Fang	sharp (75%); white (10%); pointed (5%); *long* (5%)
	Rhinestone	shiny (67%); *hard* (10%)
	Coal	black (85%); *hard* (7%)
	Rice	white (54%); small (24%); *hard* (6%)
	Hatchet	sharp (77%); metallic (5%); *hard* (5%)

inversely with associative strength of the correct response.

Task.—The experimental list consisted of 24 nouns, which constituted six sets of four nouns each in that four different stimuli in the list required the same response. The correct responses for these concepts were of high, medium, and low dominance, as described above. Degree of response dominance was derived from a standardization study by Underwood and Richardson (15). They presented nouns to 153 *Ss*, asking them to respond to each noun with one sensory adjective. The degree of response dominance (for the present list) was determined by the percentage of these standardization *Ss* giving the same response to any one noun.

Table 1 shows the 24 nouns used in this study, together with the responses given to each of these nouns by 5% or more of the standardization *Ss*. The "correct" responses or concept words are italicized. It may be seen that the concept words vary in dominance.

Interfering responses are defined as all responses given by standardization *Ss* to any one noun, other than the response designated correct. Table 1 indicates that the number and strength of interfering responses are inversely related to dominance level of the correct response; the lower the dominance of the correct response, the stronger and more numerous are the interfering tendencies.

The 24 nouns in the list were shown to *Ss* on a memory drum at a 4-sec. rate for 20 trials. There were three different successive orders of presentation, which were based on random arrangements, except that no two nouns requiring the same response were allowed to occur in succession.

Task instructions were the same for all *Ss*. Before being shown the list, *S* was told that he would see a series of 24 nouns and that he was to give a response to each noun. He was informed that there were six sets of four nouns, since the same response was correct for four different nouns. The *E* established in *S* a set to respond with sensory adjectives by asking him how he could describe objects in the simplest way, and then persisted until *S* gave several sensory descriptive adjectives. The *E* then commented, ". . . in the simplest form, we describe objects in terms of what we see, hear, smell, taste, or feel; we might give their texture, size, color, odor, shape, etc."

When *S* gave a response within the 4-sec. period, *E* told him whether the response was "Right" or "Wrong." If *S* made more than one response to a noun, which he had been asked not to do, *E* gave him information only on the first response made, and this first response was used in scoring. Failure to respond during the 4-sec. period was of course marked as a blank. Overt errors were recorded verbatim.

TABLE 2

MEAN CORRECT RESPONSES AND OVERT ERRORS IN 20 TRIALS AS A FUNCTION OF CONCEPT DOMINANCE AND ANXIETY

Anxiety Level of *Ss*	Concept Dominance						All Concepts
	High		Moderate		Low		
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean
Correct Responses							
High	147.2	17.7	136.0	23.7	98.4	32.5	381.6
Moderate	145.7	21.2	128.4	31.2	115.4	28.9	389.5
Low	153.4	8.5	133.7	27.5	118.8	26.5	405.9
Combined	148.8		132.7		110.9		
Overt Errors							
High	8.8	12.4	14.4	12.8	41.9	18.6	65.1
Moderate	9.5	14.8	19.8	13.8	31.4	15.6	60.7
Low	4.9	6.1	18.9	19.0	32.1	20.0	55.9
Combined	7.7		17.7		35.1		

Subjects.—Three groups of 31 *Ss* each were selected on the basis of their percentile rankings on the Taylor Manifest Anxiety Scale (**14**). High-anxiety *Ss* were those with percentile rankings between 100 and 80.42; moderate-anxious, between 59.01 and 40.28; and low-anxious, between 20.60 and .36.

Results

Table 2 shows the mean number of correct responses and of overt errors for the three anxiety groups at the three concept levels. Figure 1 represents the means graphically. Since Bartlett's test showed lack of homogeneity of variance for both correct responses and overt errors, analysis of variance for repeated measurements of the same *Ss* was used, which does not depend on homogeneity of variance. Analysis of variance shows that in the case of both correct responses and overt errors, there are no significant differences among groups for the entire list—that is, for all concept levels combined. However, for both correct responses and overt errors, there is a significant interaction of degree of anxiety and level of concept dominance. The interaction F for correct responses is 4.30 ($P = .01$, 4 and 180 *df*), while that for overt errors is 4.44 ($P = .01$, 4 and 180 *df*). As tested by t, there are no significant differences between groups at the high- and moderate-dominance levels. The significant differences occur at low-dominance concepts, the high-anxious *Ss* having reliably fewer correct responses than either the moderate-anxious ($t = 2.16$, $P = .05$, 60 *df*) or the low-anxious ($t = 2.70$, $P = .01$, 60 *df*). Also, the high-anxious *Ss* have significantly more overt errors than either the moderate-anxious ($t = 2.40$, $P = .02$, 60 *df*) or the low-anxious ($t = 1.99$, $P = .05$, 60 *df*).[2]

These results mean, as can be seen from Fig. 1, that the high-anxious *Ss* do as well as the other groups on the high- and moderate-dominance concepts, but their performance is disrupted on the low-dominance concepts (where interfering tendencies are strong).

The Hullian interpretation of the foregoing results involves differences in relative response strength. Before this interpretation is applied, another possibility must be eliminated—the possibility that high-anxious *Ss* have response associations which differ in kind (rather than relative strength)

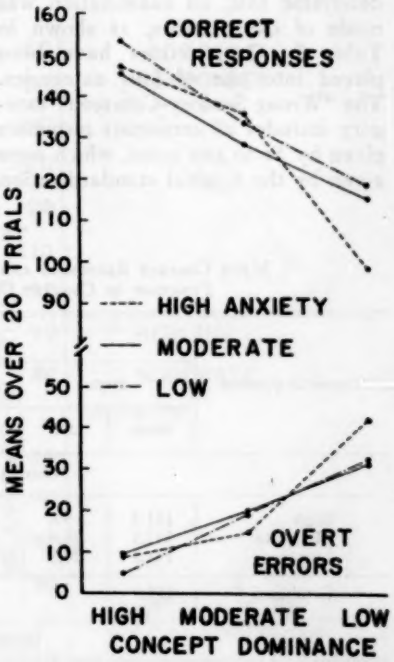

FIG. 1. Means for anxiety groups at the three concept levels.

[2] When t was obtained for groups with dissimilar variances, as shown by the F test, the table of t was entered with one-half the usual *df*. This procedure was also followed in the case of the ego-involvement groups when necessary.

TABLE 3

TYPES OF OVERT ERRORS MADE BY ANXIETY GROUPS ON LOW-DOMINANCE CONCEPTS

Degree Anxiety	Error Category								Total Errors
	Wrong Sensory: Common		Wrong Sensory: Rare		Correct for List		Nonsensory		
	No.	%	No.	%	No.	%	No.	%	
High	486	37.4	240	18.5	378	29.1	195	15.0	1299
Moderate	341	35.0	180	18.5	284	29.2	169	17.4	974
Low	387	38.9	145	14.6	235	23.6	228	22.9	995

from the response associations of groups with less anxiety. In order to determine this, an examination was made of overt errors, as shown in Table 3. Overt errors have been placed into one of four categories. The "Wrong Sensory-Common" category includes all erroneous responses given by Ss to any noun, which were given by the original standardization group with 5% or greater frequency. Thus, referring back to Table 1, it can be seen that for the noun "fang," the responses "sharp," "white," and "pointed" would be overt errors in this category. The second classification of errors includes sensory impressions which are uncommon because they were not given by 5% or more of the standardization Ss. In the third

TABLE 4

MEAN CORRECT RESPONSES AND OVERT ERRORS IN 20 TRIALS AS A FUNCTION OF CONCEPT DOMINANCE AND INVOLVEMENT

Degree Involvement	Concept Dominance						All Concepts
	High		Moderate		Low		
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean
Correct Responses							
High	151.3	9.6	130.7	28.2	102.9	38.3	384.9
Moderate	142.1	26.6	122.8	40.6	96.5	43.9	361.4
Low	144.2	20.5	130.3	23.5	85.8	32.4	360.3
Combined	145.9		127.9		95.1		
Overt Errors							
High	6.0	6.7	18.3	18.3	38.9	22.3	63.3
Moderate	9.9	14.5	20.4	22.2	36.7	22.5	67.0
Low	10.6	13.7	20.2	17.5	54.5	23.4	85.2
Combined	8.8		19.6		43.4		

category are overt errors which are correct for other words in the list, e.g., the response "round" for "rice." The fourth category includes nonsensory responses such as "animal," "good," etc.

Table 3 indicates that the differences between the three groups for low-dominance concepts are not great. Proportions for the high- and moderate-anxiety groups are nearly identical, although the highly anxious have significantly more total errors. This similarity in types of errors made suggests that the anxiety groups differ, not in the kinds of response associations but, rather, in the relative strength of associations.

Experiment II

Method

Task.—This was the same as in Exp. I. Task instructions were also the same except for the addition of ego-involving instructions.

Subjects.—Ninety volunteer *Ss* were divided into three groups of 30 each and were given different ego-involving instructions. All such instructions began, "You are probably wondering what this experiment is all about. Other studies with college students have shown that performance on this type of task. . . ."

For the low-involvement *Ss*, the completion of the instructions was " . . . is a good indicator of clerical aptitude. In other words, ability to do well here reflects the capacity to pay attention to routine details and to see their relationship to each other." For the moderate-involvement *Ss* the instructions were, " . . . is a good indicator of the degree of curiosity they possess. In other words, ability to do well here reflects curiosity about the world and how different things are related." For the high-involvement *Ss* the instructions were, " . . . is correlated highly with intelligence test scores. In other words, ability to do well here reflects the capacity to do intelligent thinking about objects in the environment and how they are related to each other."

The *Ss* were assigned to the conditions in a 1-2-3-3-2-1 sequence, in order of arrival. When the experiment was over, *Ss* were told that these instructions were for the purpose of the study, and were sworn to secrecy.

The instructions were derived from the results of a trait-ranking scale, which was based on a shorter scale developed by Hill (2). The scale was given to a total of 420 undergraduate *Ss* in two elementary psychology classes. It consisted of a list of 36 traits, e.g., adaptability, imagination, mechanical ability, perseverance, etc. The *Ss* were asked to rank the traits from 1 to 36 in order of the relative value each trait had for them. The trait with the highest mean ranking (6.52 and 6.29) in both classes was intelligence; the trait with a moderate mean rank was curiosity (20.13 and 20.03); while the trait with the lowest mean value in both classes was clerical aptitude (31.55 and 29.75).

Results

Table 4 shows the mean number of correct responses and overt errors at the three concept levels for the entire 20 trials. Figure 2 represents

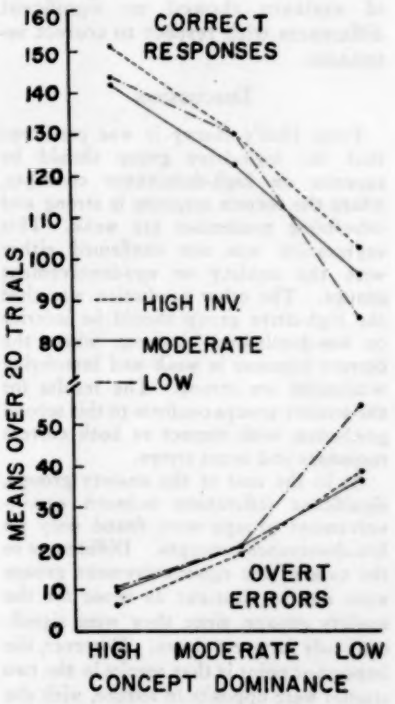

Fig. 2. Means for ego-involvement groups at the three concept levels.

the means graphically. Differences among groups are small for high- and moderate-dominance concepts, but on low-dominance concepts the low-involvement group shows fewer correct responses and more overt errors. The variances were not homogeneous, and analysis of variance for repeated measurements of the same *Ss* was used. The only significant *F* was for interaction of involvement level and concept dominance with regard to overt errors ($F = 3.63, P = .01$, 4 and 174 *df*). As indicated by *t*, low-involvement *Ss* have significantly more overt errors than either the moderate-involvement *Ss* ($t = 2.99$, $P = .01$, 58 *df*), or the high-involvement *Ss* ($t = 2.63$, $P = .01$, 58 *df*). Analysis of variance showed no significant differences with respect to correct responses.

Discussion

From Hull's theory it was predicted that the high-drive group should be superior on high-dominance concepts, where the correct response is strong and interfering tendencies are weak. This expectation was not confirmed either with the anxiety or ego-involvement groups. The other prediction was that the high-drive group should be inferior on low-dominance concepts, where the correct response is weak and interfering tendencies are strong. The results for the anxiety groups conform to this second prediction with respect to both correct responses and overt errors.

As in the case of the anxiety groups, significant differences between ego-involvement groups were found only on low-dominance concepts. Differences in the case of the ego-involvement groups were not as clear-cut as those for the anxiety groups, since they were significant only for overt errors. However, the important point is that results in the two studies were opposite in nature, with the low ego-involvement group showing poorer performance on low-dominance concepts (at least with respect to overt errors). Assuming that ego involvement as well as manifest anxiety is an indicator of drive state, results in the two studies should have been parallel. This expected parallelism was not found and only the results with the anxiety groups fit into Hullian drive theory.

A *post hoc* explanation for the findings in the ego-involvement study is that low involvement gives rise to a careless, inattentive attitude resulting in increased overt errors. However, since the expected parallelism was not found, our results indicate the need for a design in which *Ss* at each of the three anxiety levels are divided into groups receiving low-, moderate-, and high-involvement instructions.

Summary

Two separate experiments were performed: in one, three groups of *Ss* differed in level of manifest anxiety; in the other, three groups of *Ss* were given differential instructions in order to induce three different degrees of ego-involvement. All *Ss* were given a concept formation task in which associative strength of the correct response and strength of competing tendencies were varied. Predictions from Hullian theory were that a high-drive group should be superior when the correct response is strong and interfering tendencies are weak, but should be inferior when the correct response is weak and interfering tendencies strong. The study with the anxiety groups confirmed the theory to some extent since the high-anxious did more poorly than the other groups where interfering tendencies were strong. The expected parallelism between anxiety and ego involvement was not found, since the low-involvement group was inferior when interfering tendencies were strong.

References

1. Bitterman, M. E., & Holtzman, W. H. Conditioning and extinction of the galvanic skin response as a function of anxiety. *J. abnorm. soc. Psychol.*, 1952, **47**, 615–623.
2. Hill, W. F. An experimental dimensionalization of ego-involvement. Unpublished master's thesis, Northwestern Univer., 1951.
3. Hull, C. L. *Principles of behavior.* New York: Appleton-Century-Crofts, 1943.

4. Maltzman, I., Fox, J., & Morrissett, L. Some effects of manifest anxiety on mental set. *J. exp. Psychol.*, 1948, **38**, 89–101.
5. Mednick, M. T. Mediated generalization and the incubation effect as a function of manifest anxiety. Unpublished doctor's dissertation, Northwestern Univer., 1955.
6. Montague, E. K. The role of anxiety in serial rote learning. *J. exp. Psychol.*, 1953, **45**, 91–96.
7. Ramond, C. Anxiety and task as determiners of verbal performance. *J. exp. Psychol.*, 1953, **46**, 120–124.
8. Spence, K. W., & Beecroft, R. S. Differential conditioning and level of anxiety. *J. exp. Psychol.*, 1954, **48**, 399–403.
9. Spence, K. W., & Farber, I. E. Conditioning and extinction as a function of anxiety. *J. exp. Psychol.*, 1953, **45**, 116–119.
10. Spence, K. W., & Farber, I. E. The relation of anxiety to differential eyelid conditioning. *J. exp. Psychol.*, 1954, **47**, 127–134.
11. Spence, K. W., Farber, I. E., & McFann, H. H. The relation of anxiety (drive) level to performance in competitional and noncompetitional paired-associates learning. *J. exp. Psychol.*, 1956, **52**, 296–305.
12. Spence, K. W., & Taylor, J. A. Anxiety and strength of the UCS as determiners of the amount of eyelid conditioning. *J. exp. Psychol.*, 1951, **42**, 183–188.
13. Taylor, J. A. The relationship of anxiety to the conditioned eyelid response. *J. exp. Psychol.*, 1951, **41**, 81–92.
14. Taylor, J. A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, **48**, 285–290.
15. Underwood, B. J., & Richardson, J. Some verbal materials for the study of concept formation. *Psychol. Bull.*, 1956, **53**, 84–95.

(Received October 15, 1957)

Journal of Experimental Psychology
Vol. 56, No. 2, 1958

R–S LEARNING AS A FUNCTION OF MEANINGFULNESS AND DEGREE OF S–R LEARNING

ELEANORE M. JANTZ AND BENTON J. UNDERWOOD

Northwestern University

It is an established fact that, when *S* learns a list of paired associates, not only does he learn S-R (forward) associations according to instructions given him, but he also learns some R-S (backward) associations without having been instructed to do so. Thus, if A-B is the forward association being learned, B-A may be being learned at the same time. The development of such R-S associations may be inferred from transfer studies (**4, 6**), but can be shown directly by presenting *S* with B and asking him to recall A (**2, 9**).

In a previous study (**2**) it was suggested that the operations defining R-S learning may simply be another way of producing what is commonly called incidental learning. If this is true it would be expected that the variables which influence incidental learning would influence R-S learning in a comparable manner. Such comparability for a number of variables would make it likely that the different operations are producing essentially the same learning process. With regard to number of presentations and incidental learning, two studies (**1, 8**) have shown a direct relationship. Concerning meaningfulness, the data are conflicting. Brown (**1**) found no appreciable difference in the incidental learning of words and low association-value nonsense syllables. Postman, Adams, and Phillips (**7**), on the other hand, found an increase in amount of incidental learning as association value of syllables was increased from 0% to 100%. In the present study the S-R learning consists of nonsense syllables as stimuli and adjectives as responses. Since the association value of the syllables is varied, the present study will give evidence as to whether or not R-S learning is a function of the meaningfulness of the stimuli during S-R learning. In addition, the number of S-R learning trials is varied.

Since the present study was completed, a summary in English of the results obtained by a Japanese investigator, Morikawa, (**5**) has become available. In Morikawa's study both meaningfulness and number of S-R trials are reported to show a positive relationship to R-S learning. As will be seen, the present investigation confirms these results.

METHOD

General.—When R-S learning is measured by recall there is no need for a control condition. That is, if a control group, not having S-R learning, were asked to respond on an R-S recall trial it is highly unlikely that any correct responses would be given. However, if R-S learning is to be measured by transfer effects over a number of trials, a control group is essential. This control group must be at the same practice and warm-up level as the experimental group at the initiation of the transfer trials. In the present procedure three experimental groups were given 4, 12, or 24 trials on S-R learning and then 10 trials on R-S learning. Three control groups were given 4, 12, or 24 S-R trials on a comparable but different list from that given the experimental group. These groups were then given 10 trials on the same R-S list as the experimental groups. Differences between these groups on the 10 transfer trials must be attributed to R-S learning which took place as the experimental groups learned the S-R list.

Specific.—Six groups of 24 college students each, all naive to verbal-learning experiments, served as *Ss*. The *Ss* were assigned to the six groups in an ascending-descending order, i.e., to Groups, 1, 2, 3, 4, 5, 6, 6, 5, etc.

There were two lists of eight pairs each. The major list was learned by the experimental groups as their S-R list and, when "turned over," by all groups as the R-S list. Four association values were represented in the eight stimuli, there being two syllables each from 0%, 33%, 67%, and 100% values as determined by Glaze (3). The syllables and their association values are as follows: TEX, NAR (100%); POF, SUZ (67%); VOQ, KIW (33%), YUD, CIJ (0%). The eight adjectives were as follows: agog, hearty, laughing, inert, massive, equal, brutal, glassy. In S-R learning the syllables were stimuli, the adjectives were responses. Therefore, for the R-S trials the adjectives became stimuli and the syllables became responses. Four different sets of pairing of syllables and adjectives were used to minimize the possibilities of getting a pairing that might be particularly easy or particularly difficult. In actual fact the results have shown that the four different pairings were not a significant source of variance.

The S-R list given the control groups was constructed in the same manner as the one given the experimental groups except that different syllables and adjectives were used. This list was learned as the S-R list by the three control groups merely to make practice and warm-up equivalent to the experimental groups at the time the R-S list was learned by all groups. Preliminary testing showed that the two S-R lists did not differ significantly in difficulty. Similarity between the two lists was made as low as possible to minimize negative transfer when the control *Ss* shifted to the R-S list.

In S-R learning, one experimental group was given 4 trials, another 12, and the third, 24 trials. The corresponding control groups also received 4, 12, or 24 trials on their S-R list. Immediately after completing the S-R trials, all groups were given 10 trials on the R-S list. The experimental groups were instructed concerning the nature of the R-S list and were further instructed to give as many syllables as possible on the first presentation of the R-S list. They had not been told previously that they would have to learn an R-S list so there is no reason to believe that they intentionally learned R-S associations during S-R learning. The control groups were also instructed concerning the nature of the R-S list (as compared with the list that had just been presented) but were given one study trial since all items were completely new to the *Ss* in these control groups.

In presenting the lists on a Hull-type drum, the syllables were printed in capital letters, the adjectives in lower-case type. The presentation rate was 2 sec. for the stimulus and 2 sec. for the stimulus and response pair together. Three different orders of the pairs were used to minimize serial learning. A 4-sec. interval occurred between trials. Anticipation learning was used throughout, and in R-S learning the syllables were spelled.

Results

S-R learning.—The mean number of total correct responses over 4, 12, and 24 trials for the experimental groups was 5.75, 53.21, and 131.83, respectively. Corresponding values for the control groups were 6.42, 39.67, and 121.13. An analysis of variance showed that the six groups did not differ significantly in performance on the first four trials. However, considering the total correct responses over all trials, the F between experimental and control groups was significant at the 5% level. As might be expected when degree of learning is manipulated, the variance was heterogenous (Bartlett's test). The calculation of t's between the control and experimental groups for each of the three degrees of learning showed that only the two 12-trial groups differ more than would be expected by random assignment of *Ss*. The difference between these two groups was significant at the 1% level. Since the lists were shown to be not different in difficulty, the difference in performance between these two groups must represent a sampling deviation. Adjustment for these differences will be made in presenting the data on R-S learning. Meaningfulness of the stimuli was also a significant variable during S-R learning for both experimental and control groups; the higher the meaningfulness the greater the number of correct responses. However, the functions were by no means smooth, a result which might be expected since only two syllables were used as samples at each level of meaningfulness.

R-S learning measured by recall.—Immediately after learning the S-R

list the experimental *Ss* were given the R-S list in which the adjectives were now the stimuli, the syllables the responses. The *Ss* were asked to anticipate as many syllables as possible on the first presentation. Following 4, 12, and 24 S-R trials, the total number of items correct on R-S recall was 30, 79, and 95, respectively. As a function of 0%, 33%, 67%, and 100% association values, the total number of items recalled was 34, 29, 60, and 81, respectively. Very marked heterogeneity of variance was present in these recall data, but since both variables were significant far beyond the 1% level by analysis of variance it seems reasonable to conclude that both variables are positively related to R-S recall. No evidence existed for an interaction between the two variables.

It may be noted above that the syllables having 33% association values resulted in slightly lower recall than those having 0% association values. This same small difference existed in S-R learning. However, such data should not be taken to indicate a reversal of the expected relationship between association value and learning since, as noted above, two syllables from each association-value level is too small a sample to expect a precise relationship. The present data are clear in showing only that in both S-R learning and R-S recall there is a gross positive relationship between meaningfulness and performance.

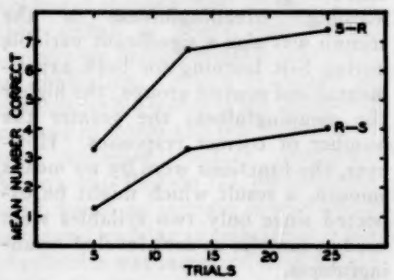

FIG. 1. Comparison of S-R and R-S learning as a function of number of S-R trials.

The increase in number of associations as a function of trials for both S-R and R-S learning is shown in Fig. 1. The R-S function simply represents the mean number of syllables correct on R-S recall. The S-R function was obtained by visual extrapolation of the S-R learning curves. For example, the acquisition curve for the 4-trial S-R group was plotted and then extrapolated to determine the approximate number of responses this group would have anticipated correctly had it been given an additional trial. The same procedure was followed for the 12-trial group to obtain an estimate of their performance on a 13th trial, and for the 24-trial group to estimate performance on a 25th trial. Since R-S recall for the three groups may be considered to have taken place on the 5th, 13th, and 25th trials, the two functions are coordinated at these three trial points. It would have been possible to use the actual performance of the 12-trial S-R group on the 5th trial to estimate the performance of the 4-trial group had they been given another trial, and the 24-trial group to estimate the 12-trial group's performance had they been given a 13th trial. However, in view of the differences in learning rate of the groups (as noted earlier) the extrapolation method would seem to be preferable and it allows comparison of S-R and R-S learning for the same *Ss* at a given point.

It is clear from Fig. 1 that the increase in number of correct associations as a function of trials is greater in S-R learning than in R-S learning. For both types of associations the greatest increase takes place during

the first half of the learning period. During the second half of the trials, S-R learning, when measured by the number of correct associations, cannot increase much more since it has almost reached the maximum of eight correct responses. This is not true for R-S learning; yet, there is only a small increase in number of responses recalled between the 13th and 25th trials. The R-S curve suggests an asymptote considerably below that of the S-R curve.

R-S learning measured by transfer.— The performance of the control and experimental groups on the ten transfer trials is shown in Fig. 2. Since the transfer list was a new list for the control groups, the first trial was a study trial. The first-trial performance of the experimental groups represents R-S recall as discussed above. An analysis of variance of total correct responses over all transfer trials allows the following conclusions: (*a*) performance of the experimental groups is superior to that of the control groups; (*b*) the higher the degree of S-R learning the greater the transfer; (*c*) the higher the meaningfulness the greater the transfer. If the performance curves for the control groups are subtracted from those for the experimental groups for each degree of learning it can be seen that the difference following 12 and 24 trials on the S-R list is maintained throughout all transfer trials. Furthermore, the difference for the 24-trial groups is greater than for the 12-trial groups. This is true in spite of the fact that earlier evidence indicated that the 12-trial experimental group was a faster-learning group and the 12-trial control group a slower-learning group than the two 24-trial groups. With

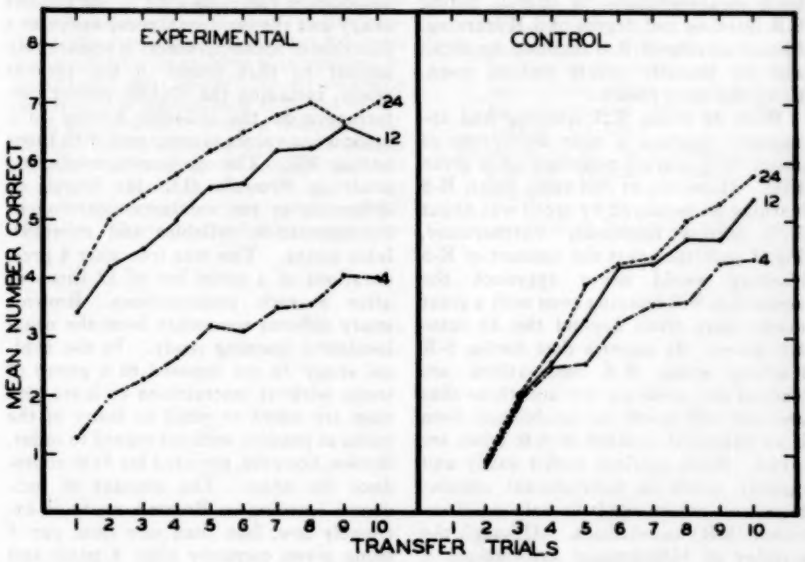

FIG. 2. Performance on each transfer trial for experimental and control groups. Amount of R-S learning is indicated by the superiority of the experimental group over the control group having had the same number of S-R trials.

only 4 trials on the S-R list the experimental group shows a clear superiority over the 4-trial control group on only the first few trials.

The need for the control groups in measuring the transfer effects is clearly apparent. Since the 12-trial control group is a slower-learning group than the 24-trial control group, these two groups may be considered roughly equivalent on these transfer trials. However, it is not possible to interpret clearly the differences in the performance of the control groups from the present data. Learning-to-learn differences, warm-up differences, and even temporary negative transfer may all influence the performance of these control groups.

Discussion

The results have shown that R-S or backward learning is a direct function of both meaningfulness of stimuli during S-R learning and degree of S-R learning. Measurements of R-S learning by recall and by transfer effects yielded essentially the same results.

With 24 trials, S-R learning had apparently reached a near asymptote of about 93% correct responses on a given trial. However, at this same point, R-S learning as measured by recall was about 50% correct responses. Furthermore, Fig. 1 indicated that the amount of R-S learning would never approach the amount of S-R learning even with a great many more trials beyond the 24 actually given. It appears that during S-R learning some R-S associations are learned and some are not and those that are not will never be established even if an extended number of S-R trials are given. Some pairings rather easily and quickly result in bidirectional associations and others result in only unidirectional (S-R) associations. Although the number of bidirectional associations is directly related to meaningfulness of the stimuli during S-R learning, the lack of evidence for interaction between degree of S-R learning and meaningfulness indicates that even with the most meaningful stimuli used here some bidirectional associations will not be formed. The situation appears analogous to results from free-association tests when it is shown that for some pairs A has a high probability of eliciting B, and B an equally high probability of eliciting A, whereas for other pairs A may have a high probability of eliciting B, but B a low probability of eliciting A.

The results of the present study would give some support for the incorporation of R-S learning into the body of data associated with what is commonly called incidental learning. The increase in R-S learning as a function of degree of S-R learning parallels the results on incidental learning as a function of number of trials (**1, 8**). The direct relationship between R-S learning and meaningfulness of stimuli during S-R learning conforms to the incidental-learning results of Postman, Adams, and Phillips (**7**). Indeed, these investigators used the same association values as used in the present study and the incidental recall curve as a function of meaningfulness is remarkably similar to that found in the present study, including the slightly poorer performance on the syllables having 33% association values as compared with those having 0%. The one known conflicting study is Brown's (**1**). He found no difference in the incidental learning of low-association syllables and monosyllabic nouns. This was true after 4 presentations of a serial list of 12 items or after 8 such presentations. Brown's study differed somewhat from the usual incidental learning study. In the typical study *S*s are exposed to a group of items without instructions to learn and then are asked to recall as many of the items as possible without regard to order. Brown, however, required his *S*s to reproduce the order. The amount of incidental learning in Brown's study is extremely low, less than one item per *S* being given correctly after 4 trials and an average of about 1.5 after 8 trials. It may well be, therefore, that differences in meaningfulness would emerge if many

more trials were given on the list than were actually given. As the data stand, however, Brown's results are contradictory to one other study on incidental learning and to the present study if R-S learning is considered incidental learning.

Summary

This study was designed to determine R-S or "backward" learning of paired associates as a function of number of S-R trials and meaningfulness of the stimuli in the S-R list. Three experimental groups were given 4, 12, or 24 trials on a list of eight paired associates having nonsense syllables for stimuli and adjectives for responses. Four different association values were represented by the stimuli. Following the specified number of trials on the S-R list, all groups were given 10 transfer trials on the R-S list which consisted of the same items as were in the S-R list but the syllables were responses and the adjectives stimuli. Three control groups had 4, 12, or 24 trials on an irrelevant list and then were given 10 trials on the same R-S list given the experimental groups.

The results show that, whether measured by recall or by positive transfer over the 10 trials, both degree of S-R learning and meaningfulness were positively related to amount of R-S learning. The data strongly suggest that R-S learning would never be complete when measured by recall even if the number of S-R trials was greatly extended over the maximum of 24 used here.

A consideration of evidence from incidental-learning studies reveals a high correspondence between the functions of meaningfulness and number of trials for incidental learning and the present results for these two variables in R-S learning. This may be used to support an argument that R-S or backward learning is essentially a form of incidental learning.

References

1. Brown, G. H. Factors influencing incidental learning. *J. exp. Psychol.*, 1954, **47**, 163–169.
2. Feldman, S. M., & Underwood, B. J. Stimulus recall following paired-associate learning. *J. exp. Psychol.*, 1957, **53**, 11–15.
3. Glaze, J. A. The association value of nonsense syllables. *J. genet. Psychol.*, 1928, **35**, 255–269.
4. Harcum, E. R. Verbal transfer of overlearned forward and backward associations. *Amer. J. Psychol.*, 1953, **66**, 622–625.
5. Morikawa, Y. Studies in paired-associate learning (1). Forward-backward recall gradient. *Jap. J. Psychol.*, 1955, **26**, 156–171.
6. Murdock, B. B., Jr. "Backward" learning in paired associates. *J. exp. Psychol.*, 1956, **51**, 213–215.
7. Postman, L., Adams, P. A., & Phillips, L. W. Studies in incidental learning: II. The effects of association value and of the method of testing. *J. exp. Psychol.*, 1955, **49**, 1–10.
8. Saltzman, I. J., & Atkinson, R. L. Comparisons of incidental and intentional learning after different numbers of stimulus presentations. *Amer. J. Psychol.*, 1954, **67**, 521–524.
9. Stoddard, G. D. An experiment in verbal learning. *J. educ. Psychol.*, 1929, **20**, 452–457.

(Received October 21, 1957)

Journal of Experimental Psychology
Vol. 56, No. 2, 1958

COMMENT ON "ROLE OF PREFEEDING IN AN APPARENT FRUSTRATION EFFECT"

ABRAM AMSEL

Newcomb College, Tulane University

A recent experiment in this Journal by Seward, Pereboom, Butler, and Jones (**8**) was designed as a test of alternative interpretations of a phenomenon reported here in 1952 by Amsel and Roussel (**3**) and termed the *frustration effect* (FE). Rats were trained in the Amsel and Roussel experiment to run out of a start box and into a runway (Runway A) to a goal box (G_1) where they found food, then to leave G_1 and run down a second runway (Runway B) to a second goal box (G_2) where they found food again. Their time to traverse Runway B was measured during a period of continuous rewards in G_1 and G_2 and, when running time had stabilized, a series of test trials were run on half of which *S* was not rewarded in G_1 prior to running in Runway B. The observation that vigor of performance in Runway B was greater following nonreward than following reward in G_1 is the FE.

The FE was interpreted by Amsel and Roussel as a motivational effect, or, more precisely, as a drive combination effect depending upon the contribution of nonreward-defined frustration to the generalized drive strength. Amsel and Roussel also recognized and considered the possibility that the test-trial difference in Runway B might be due simply to the giving or not giving of the food pellet in G_1, apart from any frustration effect—that giving a food pellet in G_1 might reduce hunger thus reducing running speed in Runway B, whereas not giving the pellet in G_1 would not reduce hunger, this failure to reduce hunger serving as the basis for faster running in Runway B after nonreward test trials. This latter interpretation they rejected for the reason that a very small amount of food (.125 gm.) was involved in the goal box preceding the test performance, and available evidence indicated that this amount of prefeeding was not sufficient to depress performance; in fact, there was some indication that prereward of this order of magnitude facilitated performance, both in terms of reduced errors in complex mazes and increased speed of running in runways (**4**, **6**).

Other experiments have demonstrated the FE with very small rewards in the first goal box of the Amsel and Roussel type of situation. For example, Roussel (**7**) employed .125 gm. of food, and Wagner (**9**) .1 gm.; Amsel and Hancock (**2**) used 2 drops of water (this has been the procedure in several other studies in our laboratory).

The Seward, Pereboom, Butler, and Jones (**8**) experiment, which is reported as a test of the alternative interpretations of the Amsel and Roussel finding, adds to the Amsel and Roussel situation a prefeeding variable: .5 or 1 gm. is given on some trials in a special chamber preceding the start box. The notion is, apparently, that if prefeeding in these amounts slows Runway B performance, there is evidence against the motivational interpretation and for the failure-to-reduce-hunger interpretation of the FE. The amounts prefed in their experiment were 1 gm. and .5 gm., 8 times and 4 times, respectively, the amount of food used on reward trials in the first goal box of the Amsel and Roussel experiment. This by itself is sufficient reason to rule out their experiment as a test of the Amsel and Roussel findings or interpretations. What they demonstrate is that prefeeding beyond the very small amounts used by Amsel and Roussel retards running in Runways A and B, which bears out the findings of earlier investigators that prereward beyond certain small amounts reverses the facilitating effect. This is brought out in their discussion where they compare running times for the four conditions of their study: *pF*—0 gm. before Runway B; *PF*—.5 or 1 gm. before Runway B; *pf*—.5 or 1 gm. before Runway B; and *Pf*—1 or 2 gm. before Runway B. They find increases in running time (decreases in vigor) from 0, to .5 or 1, to 1 or 2 gm. of prefeeding, although interestingly enough there is still clear evidence of the FE in their data. They say that this supports "the conclusion that eating a pellet before a trial retards running speed." They should have added "if the pellet weighs at least .5 gm." Amsel and Roussel were aware of this too; but the pellet they used was much smaller, and of a size which had already been shown to increase rather than decrease running speed. Although this is not a well-worked-out area, it would appear that the function relating the dependent variable, speed of running, to the independent variable, amount of prefeeding, increases from 0 gm. prefeeding to some value below .5 gm. and is then inflected downward so that, with increasing amounts of prefeeding, running speed approaches zero as a limit.

The writer takes partial responsibility for the failure of Seward et al. to use reward of appropriate magnitude. He finds, in rereading the Amsel and Roussel report, no reference to the specific size of food pellet employed. Although size of

food pellet in G_1 (.125 gm.) was reported in the later Roussel thesis, to which Seward et al. also refer, the omission in the original paper must certainly have contributed to this misunderstanding.

What puzzles the writer is that Seward et al. should have regarded their experiment as a test of the motivational hypothesis of the FE even apart from their failure to prefeed amounts equivalent to those given by Amsel and Roussel, and others, in G_1. There is, with respect to Runway B performance, a very important theoretical as well as procedural difference between prefeeding in a chamber before Runway A and feeding in G_1 following Runway A and before Runway B, regardless of whether the amount prefed is such as to increase or reduce the strength of the subsequent running behavior. Running in A is the to-be-frustrated and later the frustrated response; running in B is the frustration-motivated response. Seward et al. are in the position of reasoning that the increase in strength of running in B following nonreward in G_1 is not frustration motivated, the reasoning based on a prefeeding procedure which comes before running in A. But according to the writer's interpretation (**1**, **2**) strength of frustration resulting from nonreward in G_1 and affecting running in B depends upon strength of r_G generated in Runway A; and according to Maltzman (**5**) giving *small* amounts of prereward is one way of increasing r_G. If an appropriate amount (about .1 gm.) were to be prefed before running in A and this resulted in greater speed in A and greater FE in B than in the Amsel and Roussel situation as a control, this would fit the writer's interpretation; if a larger amount (1 gm.) is prefed and this weakens running in A and the FE compared with the Amsel and Roussel control, this still fits the same interpretation (Seward et al. claim [p. 450, par. 1] their situation produces a weaker FE than Amsel and Roussel). There is, finally, only one direct test of the FE interpretation which Amsel and Roussel rejected for their situation and which Seward et al. are unwilling to reject: to use a single runway situation comparable to Runway B and test whether rats run more slowly after eating .125 gm. in the start box than when they do not. At the risk of seeming to replay the record, the evidence strongly suggests they will not.

References

1. Amsel, A. The role of frustrative nonreward in noncontinuous reward situations. *Psychol. Bull.* 1958, **55**, 102–119.
2. Amsel, A., & Hancock, W. Motivational properties of frustration: III. Relation of frustration effect to antedating goal factors. *J. exp. Psychol.*, 1957, **53**, 126–131.
3. Amsel, A., & Roussel, J. Motivational properties of frustration: I. Effect on a running response of the addition of frustration to the motivational complex. *J. exp. Psychol.*, 1952, **43**, 363–368.
4. Bruce, R. H. An experimental investigation of the thirst drive in rats with especial reference to the goal-gradient hypothesis. *J. gen. Psychol.*, 1937, **17**, 49–60.
5. Maltzman, I. The process need. *Psychol. Rev.*, 1952, **59**, 40–48.
6. Morgan, C. T., & Fields, P. E. The effect of variable preliminary feeding upon the rat's speed of locomotion. *J. comp. Psychol.*, 1938, **26**, 331–348.
7. Roussel, J. Frustration effect as a function of repeated non-reinforcements and as a function of the consistency of reinforcement prior to the introduction of non-reinforcement. Unpublished master's thesis, Tulane Univer., 1952.
8. Seward, J. P., Pereboom, A. C., Butler, B., & Jones, R. B. The role of prefeeding in an apparent frustration effect. *J. exp. Psychol.*, 1957, **54**, 445–450.
9. Wagner, A. R. Motivational effects of non-reinforcement as a function of the reinforcement schedule. Unpublished master's thesis, State Univer. of Iowa, 1957.

(Received February 17, 1958)

Journal of Experimental Psychology
Vol. 56, No. 2, 1958

SUPPLEMENTARY REPORT: STABILITY OF A VERBAL RATE CHANGE IN EXPERIMENTAL ANXIETY[1]

FREDERICK H. KANFER

Purdue University

In a recent investigation[2] verbal rate was studied during conditioning of a tone-shock sequence. A marked increase of verbal rate and concurrent heart rate following onset of the tone was found. During the first 6 of 12 acquisition trials the verbal rate increase was obtained throughout the tone-shock interval, but at the end of training, rate increased only in the 30-sec. post-tone period. These findings suggest that with prolonged training (*a*) the effect of the signal is temporarily disruptive but disappears after additional training or is condensed to such a minimal time after tone onset that it is no longer measurable; (*b*) the effect of the "anxiety" operation may be a lasting behavioral change. The present study investigated the persistence of the effect of experimental anxiety on verbal rate over a period of prolonged training.

Method.—The *S*s were 12 male students who served at the same hour for nine consecutive days. The procedure was similar to that reported earlier, and identical on all days. The *S*s said separate words throughout each 52-min. session. All *S*s were given 17 conditioning trials under Cond. T-S (60-sec. tone co-terminal with 1-sec. shock) at varied intervals during each session.

Results.—Words were counted for each 30-sec. interval, starting with the interval preceding tone onset (Period A) and ending 30 sec. after shock (Period D), including 30-sec. post-tone (Period B) and 30-sec. pre-shock (Period C). The *F* ratio for Days (Table 1) indicates significant over-all rate changes with continued practice (increase from an over-all mean of 9.49 words/30 sec. on Day 1 to a maximal mean rate of 15.76 words/30 sec. on Day 6). Also, rates were significantly different for the four trial periods. The response rate during trials (Fig. 1) increased sharply until it reached a fairly stable level of approximately 15.5 words/30 sec. on Day 6. Period B consistently showed the highest rate. The rate difference between Periods A and B increased with training, revealing the stability of the effect of tone onset in increasing verbal rate. As training progressed, the increase was also maintained during Period C. After 4 days the rate in Period C approximated that in Period B.

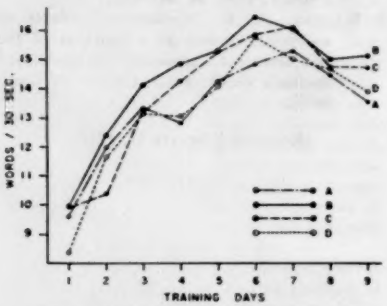

FIG. 1. Mean verbal rate for all periods of nine acquisition days.

TABLE 1

ANALYSIS OF VARIANCE OF MEAN VERBAL RATES FOR ALL PERIODS AND DAYS

Source	*df*	*MS*	*F*
Days	8	15.768	37.96*
Periods	3	2.218	5.34*
Days × Periods	24	.415	
Total	35		

* $P < .01$.

Discussion.—The increase of verbal output can be considered to represent *S*'s learning of the task. The present task appears to reach a maximum average rate of about 33 words/min., which may be asymptotic for the present procedure.

The results confirm the earlier findings that verbal rate is increased during the interval between a warning signal and subsequent shock. Prolonged acquisition training did not lead to adaptation. In fact, the effect of "anxiety" on verbal rate increased with training.

The gradual increase of rate during Period C over days suggests that *S*s formed a temporal discrimination, manifested by increased verbal rate as the shock approached. However, this discrimination did not appear to lead to a briefer and greater increase in response to the tone as was thought likely from the earlier data. With training, the effects of anxiety therefore tend to *increase* as the noxious stimulus approaches, although the early maximal response may follow shortly after onset of the signal.

(Received October 17, 1957)

[1] Supported in part by research grant M-963 from the National Institute of Mental Health, U. S. Public Health Service, while the author was at Washington University.

[2] KANFER, F. H. The effect of a warning signal preceding a noxious stimulus on verbal rate and heart rate. *J. exp. Psychol.*, 1958, **55**, 73–80.

Journal of Experimental Psychology
Vol. 56, No. 2, 1958

SUPPLEMENTARY REPORT: COMPLEXITY AND ORIENTING RESPONSES WITH LONGER EXPOSURES[1]

D. E. BERLYNE

University of California, Berkeley

In a previous experiment,[2] pairs of visual figures were exposed for 10 sec. each, and *Ss* spent more time looking at more complex and more incongruous figures. During the first 10 sec. of an exposure, *Ss* must have been occupied mainly in identifying the figures, and if the effect is confined to this initial period of identification, its theoretical significance will be limited. It may signify, for instance, merely that more complex visual stimuli take longer to identify. If, on the other hand, the preference for fixating more complex (or incongruous) figures persists long after *Ss* have had time to see what they are like, it will clearly have a more general relevance to spontaneous observing behavior, to looking "for the pleasure of looking," and to aesthetics. The present experiment was designed, therefore, to test for the effect with exposures lasting 2 min.

ment with the following exceptions: (*a*) The figures were projected on a wall 84 in. from *S*, occupying 32 × 22-in. rectangles, 48 in. apart, with lower edges 54 in. above the floor. (*b*) The second *E* (who observed eye movements) had the upper part of his body enveloped in a structure of black cloth with a cheesecloth screen in front. He made sundry sounds from time to time to disguise the coincidences between clicks coming from the recording equipment and *S*'s eye movements. (*c*) The *Ss* were told that they would be asked to estimate the time for which each pair of figures was exposed, and that noises would come from behind the screen to distract them. Questioning revealed that, thanks to these instructions and to the other expedients, the true purpose of the experiment was successfully concealed from *Ss*, which was not the case in the previous experiment.

TABLE 1

FIXATION TIMES AS PERCENTAGES OF EXPOSURE TIME

Category	Variable	10-Sec. Exposures				2-Min. Exposures			
		Less Complex Figures	More Complex Figures	Wilcoxon's		Less Complex Figures	More Complex Figures	Wilcoxon's	
				T	*p*			*T*	*p*
A	Irregularity of arrangement	37	53	53	NS	46	46	44	NS
B	Amount of material	32	51	30	<.05	33	49	24.5	<.05
C	Heterogeneity of elements	32	58	29.5	<.05	31	51	27	<.05
D	Irregularity of shape	38	53	35	NS	33	52	20.5	<.02
E	Incongruity	44	49	55.5	NS	35	58	15.5	<.01
Total		37	53	21	<.02	36	51	3	<.01

Method.[3]—The 20 pairs of figures in Categories A-E of the previous experiment were divided into four series, each containing one pair from each category in a random order. Sixteen undergraduates were divided into two groups of eight. The *Ss* in one group had a series of five pairs projected for 10 sec., followed by another series projected for 2 min. The other group had these treatments in the reverse order. The four series were used equally often in all conditions.

The procedure was as in the previous experi-

Results.—The percentages of the exposure time spent looking at the more complex (or incongruous) figures and at the less complex (or incongruous) figures in each category are shown in Table 1. The percentages for the 2-min. exposures are remarkably close to the percentages for the 10-sec. exposures (Categories A-E) in the previous experiment (i.e., 49% and 33%, respectively) and in the present experiment.

The differences for the 10-sec. exposures in the present experiment are not significant for all categories. But this is not surprising, since each *S* had five pairs of figures, as compared with 20 in the previous experiment.

(Received November 18, 1957)

[1] This study was supported by a grant from the Ford Foundation.

[2] BERLYNE, D. E. The influence of complexity and novelty in visual figures on orienting responses. *J. exp. Psychol.*, 1958, **55**, 289–296.

[3] R. Walley acted as research assistant for the experiment.

Journal of Experimental Psychology
Vol. 56, No. 2, 1958

ERRATUM

In the article "Time-Order Errors in Comparative Judgments of Hurtfulness" by Robert H. Geertsma (*J. exp. Psychol.*, 1958, **55**, 284–288), the second *P* value under Table 1 (p. 286), which is associated with the interaction of intervals and conditions, should read: $P < .10$.

JOURNAL OF EDUCATIONAL PSYCHOLOGY

This journal is now published by the American Psychological Association. In 1958 it became a bimonthly with issues in February, April, June, August, October, and December. Contents include articles on problems of teaching, learning, and the measurement of psychological development.

All back issues and subscriptions up to and including the May 1957 issue are the property of Warwick and York, Inc., 10 East Centre Street, Baltimore 2, Maryland.

Subscription $8.00
(Foreign, $8.50)

Single Copies, $1.50

Direct new subscriptions and renewals to:

AMERICAN PSYCHOLOGICAL ASSOCIATION
Publications Office
1333 Sixteenth Street, N. W.
Washington 6, D. C.

KEEP POSTED

Read

CONTEMPORARY PSYCHOLOGY

A Journal of
Reviews
Criticism
Opinion

Can you keep posted on the ever-increasing number of books now published in your field—books on social and abnormal psychology, clinical psychology, consulting psychology, psychotherapy, personality, mental health?

Let *CP* help by bringing you each month selective reviews by competent reviewers of important books, controversial books, worthwhile books, useful books.

Read the Editor's page, CP SPEAKS, for news of the publishing world, comment on communication in general and in particular, discussion of reviewing and of criticism.

Subscription, $8.00
(Foreign, $8.50)

Single copy,
$1.00

Send subscription orders to:

AMERICAN PSYCHOLOGICAL ASSOCIATION
Publications Office
1333 Sixteenth Street, N. W.
Washington 6, D. C.